



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

面向合成孔径雷达图像分类的深度学习研究方法研究

作者姓名：_____ 黄 钟 铃 _____

指导教师：_____ 雷 斌 研究员 _____

_____ 中国科学院空天信息创新研究院 _____

学位类别：_____ 工学博士 _____

学科专业：_____ 信号与信息处理 _____

培养单位：_____ 中国科学院空天信息创新研究院 _____

2020 年 9 月

A Study On Synthetic Aperture Radar Image Classification With
Deep Learning

A dissertation submitted to the
University of Chinese Academy of Sciences
in partial fulfillment of the requirement
for the degree of
Doctor of Philosophy
in Signal and Information Processing

By

Zhongling Huang

Supervisor: Professor Bin Lei

Aerospace Information Research Institute, Chinese Academy of
Sciences

September, 2020

中国科学院大学 学位论文原创性声明

本人郑重声明：所呈交的学位论文是本人在导师的指导下独立进行研究工作所取得的成果。尽我所知，除文中已经注明引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的研究成果。对论文所涉及的研究工作做出贡献的其他个人和集体，均已在文中以明确方式标明或致谢。本人完全意识到本声明的法律结果由本人承担。

作者签名：黄钟沅
日期：2020.8.20

中国科学院大学 学位论文授权使用声明

本人完全了解并同意遵守中国科学院大学有关保存和使用学位论文的规定，即中国科学院大学有权保留送交学位论文的副本，允许该论文被查阅，可以按照学术研究公开原则和保护知识产权的原则公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存、汇编本学位论文。

涉密及延迟公开的学位论文在解密或延迟期后适用本声明。

作者签名：黄钟沅	导师签名：
日期：2020.8.20	日期：

摘 要

合成孔径雷达 (Synthetic Aperture Radar, SAR) 图像分类是 SAR 图像智能解译领域中的一个基本且重要的问题, 具有广泛的实际应用价值。近年来深度学习在计算机视觉领域的飞速发展和广泛应用使其在 SAR 图像分类领域也受到了很大关注。相比于光学图像, SAR 图像分类存在很多特殊性, 包括数据上的特殊性、视觉上的特殊性、和物理散射机制上的特殊性等, 这导致 SAR 图像分类的深度学习方法研究面临以下几个主要的难点和挑战: 1) SAR 图像标记样本少, 深度神经网络容易过拟合, 泛化能力较差; 2) 视觉上与光学图像差异大, 基于视觉特征的深度模型认知困难; 3) 对 SAR 图像数据的认识不充分, 训练的神经网络缺乏对 SAR 物理散射特性的理解。目前相关的深度学习方法研究在提高 SAR 图像分类准确率上与传统方法相比获得了巨大的进展, 但仍在上述几个问题上存在一定局限性。

面向 SAR 图像分类的几个难点和挑战, 本文从 SAR 图像的视觉和非视觉信息出发, 开展了基于深度卷积神经网络的深度学习方法研究。本文的主要工作如下:

首先, 从视觉信息出发探究如何在标记样本较少的情况下利用深度卷积神经网络对 SAR 图像空间特征进行有效的学习。本文讨论了不同迁移源对 SAR 图像特征表示进行迁移学习的影响, 并根据不同的迁移源设计适用于 SAR 图像的迁移学习算法。提出了基于大规模无标签 SAR 图像的迁移学习算法, 在小样本条件下相比于当前最优 MSTAR 深度模型识别率能提升 3.17%。在迁移源与 SAR 图像相差较大的情况下提出结合域适应的传递式迁移学习算法, 改善各层特征在 SAR 图像上的迁移能力。在此基础上, 本文训练并发布了适用于 SAR 幅度图像空间特征提取的预训练网络, 成功验证了其在多种 SAR 图像分类任务中具有良好的迁移性, 在 MSTAR 任务上微调能达到 99.46% 的识别率, 优于光学预训练模型。

其次, 从非视觉信息出发利用无监督的深度学习方法探究 SAR 图像, 特别是单极化 SAR 图像复数数据中潜在的物理散射信息。基于复数 SAR 图像的时频分析理论, 提出用二维连续子带分解扩展 SAR 图像的空间二维信号。针对 SAR

图像切片提取时频信号的频域信息，用无监督手段学习频域信号中的潜在特征；针对 SAR 图像中的单个目标提取子带散射图，揭示目标在距离向和方位向连续子带的散射强度变化模式，并以此提出子带散射特性无监督分类方法，得到单极化 SAR 图像的物理散射特性结果图。

最后，本文提出融合 SAR 图像视觉和非视觉信息的深度学习分类框架。对不同场景的 SAR 图像切片，提出迁移 SAR 幅度图像预训练网络的空间特征，再融合时频信号的频域特征进行协同分类学习，改善了模型在散射复杂的 SAR 场景上的性能，与传统 CNN 相比平均准确率能提升 8.58%。另一方面，针对大型 SAR 场景中的精细地物类别理解，提出利用物理散射特性约束空间特征的半监督学习方法，仅采用极少的标记样本完成分类任务，整个学习过程相比于单纯数据驱动的神经网络更具物理解释性，并减轻了训练深度网络的数据负担。在少数训练样本条件下，加入散射特性后场景分类准确率能提升 16.58%。

综上所述，本文对 SAR 图像分类的深度学习方法进行了较深入的探究，在获得良好的 SAR 图像空间特征表示的同时，也考虑了 SAR 图像特有的物理散射特性，在 SAR 目标识别和场景分类任务上均取得了良好的效果，为今后 SAR 图像智能解译方法的发展提供了新的思路。

关键词：合成孔径雷达图像分类，深度学习，迁移学习，卷积神经网络，物理散射特性

Abstract

Synthetic Aperture Radar (SAR) image classification is an essential and important issue in SAR image interpretation field with extensive practical value. Deep learning has attracted considerable attention in SAR image classification due to the fast development and its widely successful applications in computer vision area. However, SAR image classification is a special task compared with optical data. Tackling the SAR image classification problem with deep learning is facing some difficulties that 1) limited labeled data which makes deep neural network over-fitting and lack of generalization, 2) the special visual features which result in ineffective interpretation with visually based network for complex SAR scenes, 3) the insufficient usage of SAR data which ignores the physical scattering mechanism in image classification. Although the recent researches are successful in improving the classification results compared with conventional methods, some limitations in above-mentioned aspects are still existed.

The dissertation studies the deep learning methods on SAR image classification from visual and non-visual information in SAR data. The main contributions of the dissertation are listed below:

Firstly, the effective spatial feature learning is explored with deep convolutional neural networks and transfer learning. Based on SAR target recognition, "What", "Where", and "How" to transfer are discussed to propose the effective transfer strategies. A transfer learning method based on unlabeled SAR data is proposed to obtain a robust recognition system with limited labeled SAR targets. The transitive transfer learning and domain adaptation on high-level layers are proposed to improve the transferability of features on SAR images. In order to contribute the various SAR image classification tasks, a general pre-trained model is provided and has been proved to be practical for spatial feature extraction for SAR images. It can achieve an recognition rate of 99.46% on MSTAR dataset with fine-tuning, better than ImageNet based models.

Next, the physical scattering information is explored with unsupervised deep learning methods for SAR images, especially the single-polarized SAR. The 2-dimensional

continuous sub-band decomposition based on time-frequency analysis theory is proposed to extend the spatial signals for SAR images, with non-visual information being revealed. The frequency features can be learned from a SAR image patch in an unsupervised way. Besides, a sub-band scattering pattern can be extracted for each target in SAR image, representing backscattering variations along range and azimuth directions. Then, an unsupervised learning approach is proposed to classify the scattering patterns into significant classes, and thus the scattering behavior map is obtained from a single-polarized SAR image.

Finally, two novel deep learning frameworks combining the visual and non-visual information for SAR images are proposed. On the one hand, for scene patches or target images, an end-to-end framework is proposed to fuse the spatial features transferred from pre-trained SAR image model and the frequency features of time-frequency signals, and to learn the mapping from feature space to label space. It improves the performance of distinguishing SAR image patches with complicated backscattering, performing 8.58% better than visual based CNN on overall accuracy. On the other hand, to understand the refined semantic classes of patches from one wide SAR scene, a semi-supervised framework is proposed to learn the spatial features under constraint of scattering properties. The classification task can be done with only a few labeled patches, and the scattering information can contribute to an increased accuracy of 16.58%. The novel frameworks are more interpretable than fully data-driven deep learning methods and free from the demand of a large number of labeled data for training.

To conclude, the research on SAR image classification is carried in-depth with deep learning methods. Not only learn a good representation for spatial information of the intensity image part, but also consider the specific physical scattering properties in classification. The proposed methods perform well in SAR target recognition and scene classification tasks, providing a special and promising perspective in this area.

Keywords: Synthetic Aperture Radar (SAR) Image Classification, Deep Learning, Transfer Learning, Convolutional Neural Networks (CNNs), Physical Scattering Properties

目 录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.2.1 SAR 图像分类任务的特殊性	2
1.2.2 SAR 图像分类的深度学习发展趋势	5
1.2.3 少样本下的遥感图像分类	6
1.2.4 深度学习和物理模型的结合	8
1.3 研究重点和难点	10
1.4 研究内容与主要贡献	12
1.5 论文结构安排	14
第 2 章 空间特征迁移学习	17
2.1 引言	17
2.2 特征迁移能力分析	18
2.3 迁移源分析	20
2.3.1 任务差异	21
2.3.2 数据差异	32
2.3.3 网络差异	34
2.4 传递式迁移学习	35
2.4.1 方法介绍	36
2.4.2 实验分析	36
2.5 高层特征域适应	39
2.5.1 方法介绍	39
2.5.2 实验分析	42
2.6 本章小结	45
第 3 章 幅度图像预训练模型	47
3.1 引言	47
3.2 TerraSAR-X 地物分类数据集	47
3.3 方法介绍	50
3.3.1 光学遥感数据集的传递式迁移	50
3.3.2 重采样和类别加权	51
3.3.3 联合损失函数	52
3.4 实验结果与分析	53
3.5 本章小结	57

第 4 章 面向时频特征的分类方法	59
4.1 引言	59
4.2 SAR 图像的二维连续子带分解	61
4.3 面向时频特征的分类方法	63
4.3.1 框架总览	63
4.3.2 全孔径幅度图像的迁移学习	65
4.3.3 频域信号的自编码网络	67
4.3.4 时频特征融合和后续子网络	67
4.4 实验与分析	69
4.4.1 数据集介绍	69
4.4.2 实验设置	72
4.4.3 评价指标分析	73
4.4.4 基线模型对比	75
4.4.5 可视化分析	79
4.5 本章小结	82
第 5 章 子带散射特性提取	83
5.1 引言	83
5.2 极化 SAR 图像的散射特性分类	84
5.2.1 极化 SAR 的基本概念	84
5.2.2 极化散射特性的无监督分类	86
5.3 子带散射特性聚类方法	88
5.3.1 流程总览	88
5.3.2 目标的子带散射图	89
5.3.3 多层次深度嵌入聚类	91
5.3.4 互信息评估方法	94
5.3.5 颜色编码	96
5.4 实验与分析	97
5.4.1 实验数据与实验设计	97
5.4.2 可视化与定量化分析	98
5.4.3 具体目标和场景分析	107
5.4.4 数据驱动和理论驱动方法讨论	109
5.5 本章小结	113

第 6 章 散射特性约束下的 SAR 图像分类框架	115
6.1 引言	115
6.2 散射特性约束下的分类框架	116
6.2.1 极化扩展的子带散射特性提取	118
6.2.2 散射属性提取	119
6.2.3 散射属性约束下的无监督学习	122
6.2.4 分类器设计思路	123
6.3 实际应用	124
6.3.1 高分三号 SAR 图像城区精细地物分类	124
6.3.2 哨兵一号双极化 SAR 图像海冰分类	127
6.4 本章小结	130
第 7 章 总结与展望	131
7.1 工作总结	131
7.2 研究展望	133
参考文献	135
致谢	147
作者简历及攻读学位期间发表的学术论文与研究成果	149

图形列表

1.1 SAR 图像分类任务的特殊性·····	3
1.2 SAR 图像分类深度学习方法的发展趋势 ·····	6
1.3 深度学习和物理模型的关系 (Karpatne 等, 2017) ·····	10
1.4 本文的研究重难点分析 ·····	11
1.5 本文的章节组织结构·····	14
2.1 源任务 S^i 第 k 层特征迁移能力分析示意图 ·····	19
2.2 源任务 S^i 和源任务 S^j 的第 k 层特征迁移能力对比分析示意图 ·····	19
2.3 本实验所用 OpenSARShip 数据集的三类目标示意图 ·····	20
2.4 基于无标签 SAR 图像的 SAR 目标识别迁移学习网络架构 ·····	21
2.5 卷积自编码器单元结构示意图 ·····	22
2.6 栈式卷积自编码器堆叠示意图 ·····	25
2.7 CNN-baseline 和 CNN-TL 训练集和测试集损失函数变化情况。(a) CNN-baseline, (b) CNN-TL. ·····	27
2.8 CNN-TL 模型加入反馈旁路前后测试集损失函数的变化情况 ·····	28
2.9 训练样本减少情况下 MSTAR 目标识别结果 (%) ·····	28
2.10 无标签 SAR 场景图像示例 ·····	29
2.11 SAR 场景图像在堆栈自编码器各层训练过程中的重建结果 ·····	29
2.12 重建任务和分类任务迁移时第一层卷积核的变化情况 ·····	30
2.13 不同源任务迁移到 OpenSARShip 目标识别任务上时网络各层特征的迁移性变化 ·····	31
2.14 三种迁移数据源 ·····	32
2.15 不同迁移数据源在 AlexNet_Conv 网络上预训练后每层特征迁移到 OpenSARShip 目标识别上的表现 ·····	33
2.16 MSTAR 迁移源数据在不同网络上进行预训练后每层特征迁移到 OpenSARShip 目标识别上的表现 ·····	35
2.17 传递式迁移学习示意图 ·····	37
2.18 传递式迁移学习前后 H_Net 的每层特征迁移到 OpenSARShip 目标识别上的表现 ·····	38
2.19 传递式迁移学习前后 AlexNet_Conv 的每层特征迁移到 OpenSARShip 目标识别上的表现 ·····	38
2.20 一体式域适应迁移框架 (ITL) 和独立式域适应迁移框架 (STL) 示意图。(a) ITL, (b) STL. ·····	42

2.21	MSTAR 和 OpenSARShip 数据在不同预训练模型的第四层和第五层特征的分布情况	43
2.22	MSTAR 和 OpenSARShip 在 AlexNet_Conv 两种预训练模型第五层的特征域适应前后的可视化效果	44
3.1	TerraSAR-X 大规模 SAR 地物标记数据集统计信息	48
3.2	数据集地物多样性示例	49
3.3	噪声标签示例	50
3.4	ResNet-18 主干网络示意图	50
3.5	TerraSAR-X 数据集与 NWPU-RESISC45 数据集 (Cheng 等, 2017) 各类别标签对比与图像示例	51
3.6	三种训练方式下停止时间、损失函数变化以及测试准确率对比	54
3.7	测试集上的各损失函数变化曲线	54
3.8	预训练模型在测试集二 7 个集合中分类准确率	56
3.9	模型在测试集一和测试集二上的混淆矩阵和 Top-2 准确率结果。(a) 测试集一, (b) 测试集二。	56
3.10	预训练模型利用类别激活映射 (CAM) (Zhou 等, 2016a) 对一景 SAR 图像中几种地物的可视化热度图。(a) 居所, (b) 工业区, (3) 水体。 ..	57
3.11	预训练模型迁移至 MSTAR 目标识别任务的测试结果	57
4.1	SAR 图像目视解译难点示例。(a) Sentinel-1 卫星对同一个地区的成像对比, (b) 居民区和集装箱的 SAR 图像示例, (c) 船只和工业设施的 SAR 图像示例。	60
4.2	结合散射特性的 SAR 图像分类问题描述	61
4.3	SAR 时频四维信号的理解	63
4.4	DSN 框架概述	64
4.5	单视复数图像数据的强度图像和多视地距图像对比示意图	65
4.6	DSN 具体实现	66
4.7	残差模块	66
4.8	S1 数据集	70
4.9	TSX 数据集	71
4.10	OPS 数据集	72
4.11	DSN 和 CNN 在 S1 数据集上的表现。(a) F1 分数, (b) 平均准确率。	74
4.12	DSN 和 CNN 在 TSX 数据集上的表现。(a) F1 分数, (b) 平均准确率。	74
4.13	DSN 和 CNN 在 OPS 数据集上的表现。(a) F1 分数, (b) 平均准确率。	74
4.14	S1 数据集用 90% 的数据进行训练时, 每类地物在 DSN 和 CNN 模型中的准确率和 F1 分数。(a)F1 分数, (b) 准确率。	75

4.15 DSN 和 CNN 模型的 ROC 曲线。(a) S1 数据集自然地物类别, (b) S1 数据集人造目标类别, (c) TSX 数据集, (d) OPS 数据集。·····	76
4.16 DSN 与基线模型的平均准确率对比。(a) CV-CNN 和 DSN, (b) CNN、F-CNN 和 DSN。·····	77
4.17 S1 数据集在 DSN 和 CNN 中的特征分布。(a) DSN, (b) CNN。·····	79
4.18 自然地物样本频域信号幅度图可视化·····	79
4.19 图 4.17(b) 中所示 A、B、C 样本在 DSN 和 CNN 中的预测结果对比及频域信号分析。(a) 示例 A, (b) 示例 B, (c) 示例 C。·····	81
4.20 TSX 数据集在 DSN 和 CNN 中的特征分布。(a) DSN, (b) CNN。·····	82
5.1 基于时频分析的无监督子带散射特性聚类方法主要流程图·····	88
5.2 提取目标子带散射图的时频分析方法·····	89
5.3 几种典型目标和非典型目标的子带散射图·····	91
5.4 深度卷积嵌入聚类方法示意图·····	92
5.5 卷积自编码网络示意图·····	92
5.6 高分三号全极化实验数据。(a) 巴黎, (b) 旧金山。·····	97
5.7 对于 512×512 大小的 SAR 图像切片选取不同 k 值时 DEC 的训练结果	99
5.8 巴黎地区。(a) 全极化 SAR 的极化散射特性分类结果图, (b) 单极化 SAR 的子带散射特性分类结果图。·····	100
5.9 旧金山地区。(a) 全极化 SAR 的极化散射特性分类结果图, (b) 单极化 SAR 的子带散射特性分类结果图。·····	101
5.10 互信息矩阵 (巴黎地区)。(a) 第一层训练, (b) 第二层训练。·····	103
5.11 子带散射类别 $a3$, $c3$, $e3$ 和 $g1$ 的子带散射图示例·····	104
5.12 互信息矩阵 (旧金山地区)。(a) 第一层训练, (b) 第二层训练。·····	105
5.13 旧金山地区极化散射特性 $M2$ 和子带散射类别 f 的结果可视化图·····	106
5.14 旧金山地区两个目标的子带散射类别 f 和极化散射特性 $O2$ 及 $D1$ 的结果可视化图·····	106
5.15 旧金山地区人造目标散射特性结果分析·····	107
5.16 巴黎埃菲尔铁塔散射特性结果分析·····	108
5.17 巴黎地区人造目标散射特性结果分析·····	108
5.18 巴黎地区自然地物散射特性百分比分析·····	109
5.19 旧金山地区自然地物散射特性百分比分析·····	110
5.20 将 MODEL-A 应用到不同 SAR 场景中的实验结果对比·····	111
5.21 MODEL-A 测试巴黎地区大小为 6144×7680 的全景 SAR 图像的结果。 (a)HH 通道 SAR 图像, 左下角的红框标记为训练数据, (b)MODEL-A 的测试结果。·····	112
5.22 MODEL-P 在 Sentinel-1 SAR 图像上测试和微调的结果·····	113

6.1 SAR 图像的认知过程	115
6.2 散射特性约束下的 SAR 图像分类框架	117
6.3 对不同极化方式的极化子带散射图输入。	120
6.4 不同极化方式下的（极化扩展）子带散射特性分类结果示意图	124
6.5 训练样本和测试样本示意图	125
6.6 散射主题提取结果示意图	126
6.7 格陵兰岛所选实验数据示意图	128
6.8 H/α -Wishart(Cloude 和 Pottier, 1997) 分类结果	129
6.9 SAR 图像切片和对应散射特性图示例	129

表格列表

1.1 三大研究内容在对应章节、待解决问题、解决思路和相关应用层面上的总结	12
2.1 本实验所用的 OpenSARShip 数据集子集	20
2.2 分类支路各层网络设计	22
2.3 MSTAR 数据集	26
2.4 SAR 目标识别算法在 MSTAR 数据集上的测试结果	26
2.5 CNN-baseline 模型 MSTAR 目标分类结果混淆矩阵	27
2.6 CNN-TL 模型 MSTAR 目标分类结果混淆矩阵	28
2.7 目标重建任务迁移过程中每层卷积核 L2 范数的变化百分比	30
2.8 目标分类任务迁移过程中每层卷积核的 L2 范数变化情况	30
2.9 三种深度卷积神经网络的结构设计及在 MSTAR 和 OpenSARShip 数据集上从头训练的识别率	35
2.10 迁移源分析总结	36
2.11 采用不同网络和源数据冻结各层特征进行 OpenSARShip 船舶目标识别的迁移时的识别率	39
2.12 ITL 和 STL 两种算法相比于微调的性能提升	45
3.1 TerraSAR-X 数据集训练集和测试集划分	53
3.2 在测试集一和测试集二上的测试结果	55
4.1 DSN 框架三个子网络的卷积核设计	68
4.2 Sentinel-1 条带模式下的 SLC 图像参数	70
4.3 TerraSAR-X 卫星聚束模式下的 SSC 图像参数	71
4.4 S1 数据集在 DSN、TL-CNN 和 CNN 中各类别测试集的表现 (准确率%)	77
4.5 S1 数据集在 DSN、TL-CNN 和 CNN 中各类别测试集的表现 (F1 分数)	78
5.1 高分三号实验数据信息	98
5.2 训练源数据与应用的目标 SAR 数据各项参数对比	112
6.1 Top-1 和 Top-2 测试准确率	127
6.2 哨兵一号海冰分类数据详情	128
6.3 哨兵一号 SAR 图像海冰分类实验结果	130

缩略语列表

BoST	Bag of Scattering Topics
BoW	Bag of Words
CAM	Class Activation Map
CNN	Convolutional Neural Network
DEC	Deep Embedding Clustering
DLR	Deutsches Zentrum für Luft- und Raumfahrt
DSN	Deep SAR-Net
FN	False Negative
FP	False Positive
GRD	Grounded Range Detected
HDEC-TFA	Hierarchical Deep Embedding Clustering based on Time-Frequency Analysis
ITL	Integrated Transfer Learning
IW	Interferometric Wide Swath
LDA	Latent Dirichlet Allocation
LULC	Land Use and Land Cover
MGD	Multi-looked Grounded Detected
MMD	Maximum Mean Discrepancy
OA	Overall Accuracy
PMI	Point-wise Mutual Information
PolSAR	Polarimetric SAR
ResBlk	Residual Block
ResBnkBlk	Residual Bottleneck Block
RKHS	Reproducing Kernel Hilbert Space
ROC	Receiver Operating Characteristic

SAR	Synthetic Aperture Radar
SCAE	Stacked Convolutional Auto-Encoder
SL	SpotLight Mode
SLC	Single Look Complex
SSC	Single-looked Slant Complex
SVM	Support Vector Machine
STL	Separated Transfer Learning
TFA	Time-Frequency Analysis
TP	True Positive
VHR	Very High Resolution

第1章 绪论

1.1 研究背景及意义

合成孔径雷达 (Synthetic Aperture Radar, SAR) 是一种主动式微波传感器, 可以不受光照和气候的影响进行全天时全天候的成像, 在军事和民用的对地观测领域均有广泛应用。从上世纪七十年代美国国家航空航天局 (NASA) 发射的第一颗装载 SAR 的 SeaSat 卫星起, SAR 系统技术经历了近半个世纪的迅速发展, 从一开始以单极化、单一成像模式为主的 SAR 系统——例如上世纪九十年代欧空局 (ESA) 发射的 ERS-1/2 和日本的 JERS-1 卫星, 到现在多种极化方式、多种成像模式的高分辨率 SAR 系统——例如德国的 TerraSAR-X/TanDEM-X 系统、我国的高分系列卫星等, 分辨率已能从最初的几十米发展到米级甚至达到亚米级。如今 SAR 数据的获取能力不断提升, 数据量级急剧增加, 依靠人工目视解译时效性太低, 在海量数据中迅速挖掘有用的信息应用到资源勘探、军事侦察、海洋探测等多个领域中显得尤为重要, 发展高效的 SAR 图像智能解译技术迫在眉睫。

SAR 图像分类作为 SAR 智能解译领域中的一个基本问题, 长期以来受到广泛关注, 包括目标识别、土地使用和土地覆盖类别划分、SAR 图像地物分割等应用。近年来深度学习在计算机视觉领域取得了巨大成功, 许多研究也将其应用到 SAR 图像分类任务中, 一般采用端对端的神经网络模型对 SAR 图像进行特征提取和分类, 自动学习层级特征, 克服手动设计特征和分类器带来的局限性。深度学习方法为 SAR 图像的智能解译注入了新的活力, 但由于 SAR 图像不同于光学自然图像的成像机理, 深度学习方法在 SAR 图像分类中的应用也存在很多特殊性。

SAR 作为一种主动微波成像系统, 发射一定频率的电磁波在空间中传输, 地表环境和地面目标对入射电磁波产生调制效应后, 散射的电磁波被雷达接收形成回波信号, 进行一系列的成像处理后生成二维图像。SAR 接收的电磁波包含了多种信息, 例如幅度信息、相位信息、频率信息和极化信息等, 反映了地面目标对电磁波的散射特性, 与目标的结构、方位、材质, 以及 SAR 的工作波段、入射角等因素都紧密相关。因此对于 SAR 图像, 不仅包含了人眼可见的“图像”, 也包含了人眼不可见的“信号”。一方面从视觉上, SAR 图像具有相干斑噪声、

透视收缩、顶底倒置等特殊的视觉效果，并且不具备光学图像中目标常具有的旋转不变性等特点，目标散射中心强度和分布对方位角极其敏感，习惯了光学视觉系统的人类尚且理解困难，由此导致人工标注代价高，缺少大量的训练样本。另一方面，SAR 图像中隐含的雷达特性提供了丰富的物理散射信息，反映了目标的大小、几何结构、材料属性等物理特性，是进行 SAR 图像特征提取和分类的重要依据。

SAR 图像分类有别于普通的自然图像或光学遥感图像的分类研究，根据 SAR 图像分类应用中存在的特殊性分析其中关键性的难点，开展有特色的适用于 SAR 图像的深度学习新框架和新方法，对 SAR 图像智能解译领域的发展具有重要意义。

1.2 国内外研究现状

1.2.1 SAR 图像分类任务的特殊性

SAR 图像的分类任务可以粗略分为目标和场景两大块，对于 SAR 目标而言一般以识别和检测为主，其中检测可视作特殊的分类任务；SAR 图像的场景分类问题又可以划分为两个部分，一是面向一定尺寸的 SAR 图像切片，将其划分为多种土地使用和土地覆盖类别 (Land-Use and Land-Cover, LULC)，根据 SAR 图像切片中的目标和地表形态来区分功能性的土地使用类别和形态性的土地覆盖类别；另外是面向 SAR 图像进行逐像素的地物类别划分，一般针对极化 SAR 图像进行。下面将依次阐述这三种 SAR 图像分类任务的特殊性。

与传统计算机视觉领域中面向自然图像的研究相比，遥感领域的目标识别和检测任务具有一定的特殊性。例如遥感目标的检测面临背景复杂、目标尺寸过小、排列紧密有遮挡等挑战 (Yang 等, 2019)，其中因为观测角度不同造成遥感目标的旋转多样性被认为是最重要的问题之一 (Cheng 等, 2016)；而遥感目标的识别通常需要做到型号级的认知，例如区分不同类型的船只 (Hou 等, 2019)、飞机 (Liu 等, 2013) 或坦克 (Dong 等, 2014) 等，需要充分考虑遥感目标的细节和特征。对 SAR 目标的研究更具有特殊性，例如在 SAR 图像中存在透视收缩、阴影等现象，Papson 和 Narayanan (2012) 提出利用 SAR 目标的阴影中隐含的信息辅助目标识别。图1.1中第一行给出了几个实例，如目标在方位向上与雷达平台的相对运动导致的多普勒模糊，舰船上强散射点在方位向上产生的线装散焦，以及电磁

波的多径散射在目标附近产生的虚像，都会对 SAR 目标的检测和识别带来难度 (Wang 等, 2016b); 另外 SAR 目标对方位角及其敏感, 雷达在不同角度对同类别目标的成像结果相差很大, 在实际应用中 SAR 目标样本有限, 一般不可能有大量不同角度下的成像数据用于训练识别系统 (Pei 等, 2018a,b)。

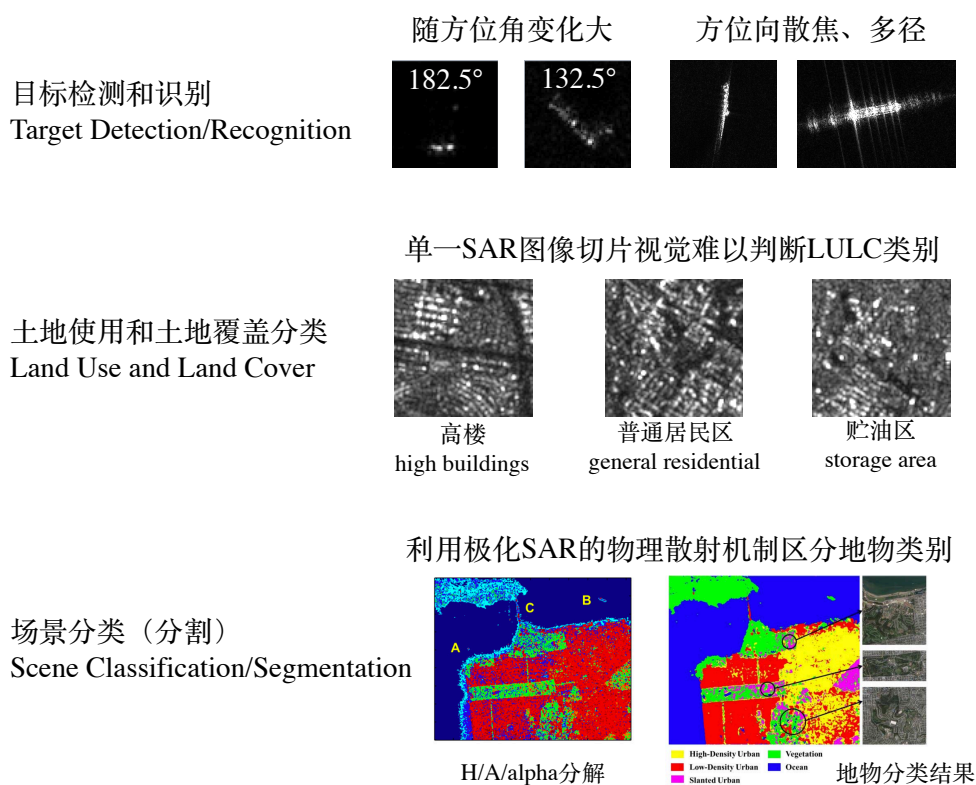


图 1.1 SAR 图像分类任务的特殊性

Figure 1.1 The characteristics of SAR image classification tasks.

基于切片的光学遥感图像 LULC 分类目前有 UC-Merced(Yang 和 Newsam, 2010)、AID(Xia 等, 2017)、GID(Tong 等, 2020)、NWPU-RESISC45(Cheng 等, 2017)、BigEarthNet(Sumbul 等, 2019) 等公开数据集, 涵盖航空遥感图像、Google Earth 卫星遥感图像、Sentinel-2 多光谱遥感图像等数据, 包括几十种 LULC 类别, 其中 BigEarthNet(Sumbul 等, 2019) 数据集共有 590,326 张多波段的 Sentinel-2 卫星遥感图像切片, 并且赋予了多类别的标签, 为遥感图像 LULC 分类提供了丰富的基准数据。相比之下, SAR 图像不符合人类的视觉认知, 数据标注的成本较高, 操作更困难, 目前对于大规模的 SAR 图像 LULC 分类任务还缺乏较多的公开标注样本, Dumitru 等 (2016) 基于 TerraSAR-X 的条带模式 SAR 图像数据提出一个 LULC 分类的数据集但未公开, 近日 Zhao 等 (2020) 基于 Sentinel-1 卫星

的 SAR 数据提出 OpenSARUrban 数据集，先利用光学遥感图像进行标注之后再对应到 SAR 图像中。通过可视化这两个数据集中的样本会发现，只依靠单一的 SAR 图像切片，人眼很难分辨出其所属类别，特别是与人类活动密切相关的土地使用类别，如图1.1的第二行给出的 OpenSARUrban 数据集中高楼、普通居民区、贮油区三个土地使用类别的 SAR 图像切片示例。

在 SAR 图像逐像素的场景分类问题中，相干斑会使像素点的分类造成噪声，需要进一步的后续处理，例如Geng 等 (2015) 和Geng 等 (2017) 的工作都是围绕单极化 SAR 图像的场景分割问题进行，在神经网络中设置了平均池化的尺度变换层来平滑掉相干斑噪声，在后处理中采用图割来抑制相干斑对分割结果的影响。更多情况下，SAR 图像逐像素的场景分类应用于极化 SAR 图像中，因为目标极化矩阵包含了丰富的极化信息，反映了目标的物理散射特性，通过目标极化分解方法来描述目标的物理散射机制，从而对应不同的地物类别。例如Cloude 和 Pottier (1997) 提出的 $H/A/\alpha$ 目标极化分解方法将散射特性映射到 $H-\alpha$ 平面的九个区域中，不同区域对应水体、城区、不同类型的植被等地物；Freeman 和 Durden (1998) 提出 Freeman-Durden 三分量分解方法，将偶极子散射、二次散射和布拉格散射进行建模。Jiao 等 (2014) 利用 Cloude-Pottier(Cloude 和 Pottier, 1997) 和 Freeman-Durden(Freeman 和 Durden, 1998) 分解用于农作物的分类。极化 SAR 图像的地物分类虽然从结果上来看，与自然图像或光学遥感图像中的场景分割类似，都是逐像素的分类任务，但光学图像的场景分割主要还是基于视觉上的特征以及像素间的关系，而 SAR 图像的场景分类主要基于目标的极化散射特性。这也说明理解 SAR 图像的物理散射特性对 SAR 图像的分类具有重要的意义。

SAR 图像在数据上也具有特殊性。自然图像或普通光学图像一般为单通道（灰度图像）或 RGB 三通道（彩色图像）的实数矩阵，而 SAR 图像由于产品级别和成像模式的差别具有多种形式。斜距单视复数图像级别的产品一般以 16 位或 32 位的复数数据构成，具有实部和虚部，同时包含幅度和相位信息，在经过定标、地距投影、辐射校正、多视、量化等处理之后，得到去相干斑噪声、只保留幅度信息的单通道 16 比特实数多视地距图像，在 TerraSAR-X、Sentinel 1、高分三号中分别对应 MGD、GRD、L1B 格式的数据，一般情况下基于深度学习的 SAR 图像分类方法将只保留幅度信息的 SAR 图像数据作为神经网络的输入(Wurm 等, 2019; Kang 等, 2017; An 等, 2018)。但在这一系列处理过程中丢失了相

位信息,从某种程度上造成了信号的损失,不利于充分利用 SAR 图像的信息进行判别。极化 SAR 的复数数据中包含极化信息,干涉 SAR 通过复数图像对的共轭相乘得到的干涉图相位值对提取三维信息尤为重要,因此在进行 SAR 图像分类的深度学习研究方法研究时还需考虑如何正确应用复数数据。

基于以上的调研分析可以看出, SAR 图像分类任务与光学图像的分类相比有很大的差异,如何根据这些特殊性去设计深度学习算法来解决 SAR 图像分类问题十分重要。

1.2.2 SAR 图像分类的深度学习方法发展趋势

SAR 图像分类中的深度学习方法正在经历三个阶段的发展,如图1.2所示。

第一阶段的研究主要集中在首次将深度神经网络引入 SAR 目标识别、目标检测、SAR 场景分类、分割等任务中,用数据驱动的深度网络学习自动特征提取,以替代传统方法中的人工设计特征,各类评价指标相比于传统方法得到了明显的提升。这个时期的代表性研究包括迁移已有的自然图像预训练网络进行微调 (Kang 和 He, 2016; Wang 等, 2017); 用自编码器 (Auto-Encoders, AE) (Xie 等, 2014; Geng 等, 2015; Hou 等, 2016; Zhang 等, 2016)、受限玻尔兹曼机 (Restricted Boltzmann Machine, RBM) (Qin 等, 2017)、深度信念网络 (Deep Belief Network, DBN) (Lv 等, 2015; Zhao 等, 2017c) 等无监督学习方法自动学习 SAR 图像的多级特征,再训练支撑向量机 (Support Vector Machine, SVM) 或 k 最近邻 (k Nearest Neighbor, kNN) 等分类器。这个阶段的研究证明了深度神经网络通过自动学习得到的多层次特征效果优于传统方法中常用的特征,如傅里叶描述子 (Anagnostopoulos, 2009)、基于主成分分析 (Principal Component Analysis, PCA) 的特征 (Mehra 等, 1998)、多尺度的尺寸不变特征变换 (Scale Invariant Feature Transform, SIFT) 算子 (Ruan 等, 2016)、单演信号 (Dong 等, 2014) 等。

第二阶段的研究主要针对特定的 SAR 图像分类任务设计专门的端对端深度神经网络从头训练,改进网络设计方案,优化网络结构和损失函数,使之更符合特定的 SAR 任务以取得最优的效果。例如Chen 等 (2016) 设计了一个小型的全卷积网络用于 MSTAR 目标数据集的识别,首次取得优于 99% 的识别率; Lin 等 (2017) 提出将深度卷积高速单元 (Highway Unit) 用于少量 SAR 目标的自动识别; Minh 等 (2018); Ndikumana 等 (2018) 利用循环神经网络 (RNN) 和长短时记忆模块 (LSTM) 对 Sentinel-1 的时间序列数据进行农作物的研究; Geng 等 (2017) 针

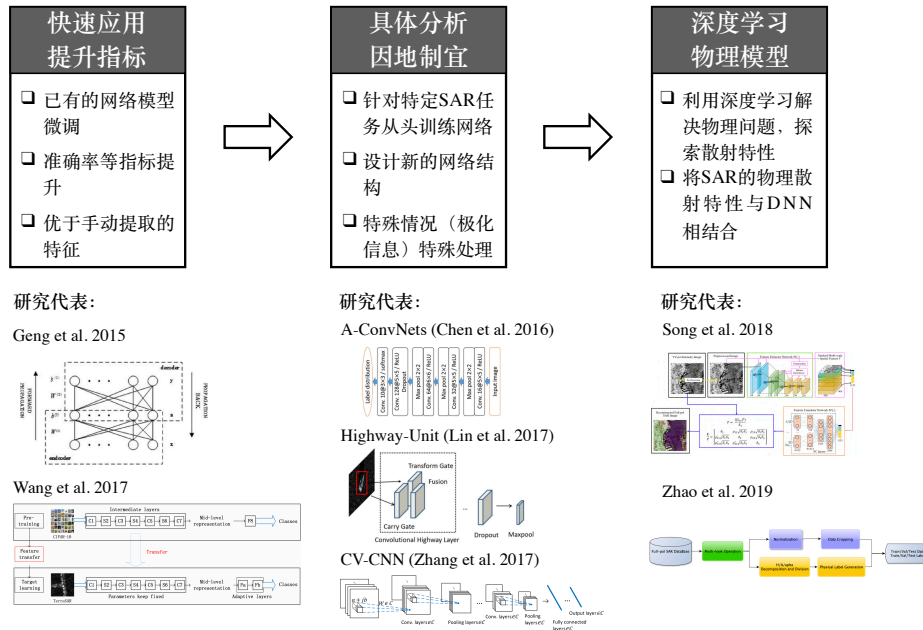


图 1.2 SAR 图像分类深度学习方法的发展趋势

Figure 1.2 The development trend of deep learning methods in SAR image classification.

对 SAR 图像相干斑噪声的影响提出了深度监督收缩卷积神经网络用于 SAR 图像的分割；Zhang 等 (2017) 提出了复数卷积神经网络用于极化 SAR 图像的分类；Jiao 和 Liu (2016) 针对极化 SAR 分类设计了一种 Wishart 深度堆叠网络，将快速计算的 Wishart 距离作为网络自动分类时的相似性度量等。

第三阶段的研究则注重在深度学习方法中结合已有的专业知识和 SAR 图像反映的物理散射特性，融合数据驱动的深度模型和理论驱动的物理模型进行协同分类。一方面运用深度学习可以在大量数据中自动挖掘未知潜在模式的优势，在缺乏可靠物理模型支撑的情况下（例如缺少极化信息）开展揭示物理散射特性的研究 (Song 等, 2018; Zhao 等, 2019)；另一方面在 SAR 图像分类的深度学习框架中加入散射信息进行干预，使深度网络模型在理解复杂 SAR 场景时给出具有物理解释性的结果。这一阶段的研究目前仍处于起步阶段，在 1.2.4 节将会给出详细介绍与说明。

1.2.3 少样本下的遥感图像分类

在遥感图像分类的深度学习方法研究中会经常面临标记样本较少的困境，特别是 SAR 遥感图像，为训练深度神经网络带来了极大的难度，容易造成过拟合导致模型泛化能力较差。针对这个问题目前主要有数据扩充和模型优化两种类

型的策略,其中数据扩充主要目的是扩大训练样本的规模帮助模型训练,模型优化则通过改善网络结构避免过参数化或采用迁移学习等有效策略防止过拟合(潘宗序等,2019)。

在大多数的遥感图像深度学习分类方法中,都对图像采取了平移、翻转、加噪声等基本的数据扩充手段。对于更特殊的 SAR 图像,Ding 等(2016)针对 SAR 目标对方位角比较敏感的特性,提出基于方位角的样本合成方法,用于生成特定方位角的 SAR 目标样本,扩大训练样本的规模,Malmgren-Hansen 等(2017)通过仿真 SAR 目标来补充样本数量,Cui 等(2019); Bao 等(2019)提出基于生成对抗网络学习的样本生成方法。

上述提出的第二阶段的研究中有一些代表性的工作侧重于模型优化方面,根据所持数据的规模设计小型的神经网络防止过参数化。例如面向 MSTAR 数据集的 SAR 目标识别任务中,Chen 等(2016)提出的 5 层全卷积神经网络取得了 99.13% 的识别率;Wang 等(2016a)设计了两个卷积层和一个全连接层相连接的卷积网络用于双极化 SAR 图像的冰川浓度估计;一些浅层的堆叠自编码器或玻尔兹曼机被用于极化 SAR 图像的分类(Zhang 等,2016;Lv 等,2015),并且Lv 等(2015)发现当隐藏层数大于 2 之后,由于过拟合问题,网络的分类性能反而下降。

另一个常用的优化策略则是采用迁移学习的方法,这在第一阶段的研究中是最直接有效的方式,为解决训练数据缺乏的问题提供了有效的途径。例如在光学遥感图像的应用中,很多研究都采用自然图像的预训练网络进行迁移。Hu 等(2015)首先将不同的 ImageNet 预训练模型,例如 CaffeNet、AlexNet 和 VGG-16 迁移到三波段的光学遥感图像分类任务中,取得了很好的效果。很多后续的研究也采用了类似的手段,将 ImageNet 预训练模型迁移到不同的遥感任务中,例如图像配准(Wang 等,2018b)、飞机检测(Chen 等,2018)、场景分类(Marmanis 等,2016;Zhao 等,2017a)、遥感图像分割(Fu 等,2017)以及超分辨重建(Yuan 等,2017),分别涉及到高光谱、多光谱遥感图像等。由于光学遥感图像呈现出地物的光波反射特性,与自然图像在视觉上具有很大的相似性,因此在光学遥感图像的应用中,从自然图像的预训练网络进行迁移的方法取得了一系列的成功。除了自然图像,也有研究采用了其他的迁移数据源,例如Windrim 等(2018)提出在不同平台的传感器的高光谱成像数据之间进行迁移,相似的还有Samat 等(2016)

的研究。

由于 SAR 图像反映了地物的微波散射特性而非光波反射特性,呈现出众多的非视觉特性,目前对于光学图像是否能有效迁移到 SAR 图像上还具有一定的争议。[Kang 和 He \(2016\)](#) 利用在 CIFAR-10 数据集上训练得到的网络的中间层特征应用到 TerraSAR-X 图像中;[Wang 等 \(2017\)](#); [Zhong 等 \(2019\)](#); [Wang 等 \(2019\)](#); [Lu 和 Li \(2019\)](#) 选择直接用单通道的 SAR 图像数据在 ImageNet 的预训练网络上进行微调, [Wu 等 \(2018\)](#) 在极化 SAR 的场景分类中采用了 VGG-16([Simonyan 和 Zisserman, 2015](#)), ResNet-50([He 等, 2016](#)) 和 DenseNet-121([Huang 等, 2017a](#)) 在 ImageNet([Deng 等, 2009](#)) 上的预训练网络,用于极化 SAR 图像的空间特征提取。在反对的声音中,认为由于 SAR 与光学遥感图像成像机制的不同,二者在视觉上呈现出很大的差别,因此基于光学图像训练的模型包括 ImageNet 和光学遥感图像,对 SAR 图像并不具备良好的迁移性,甚至会削弱深度神经网络对 SAR 图像的特征提取效果([Marmanis 等, 2017](#); [Wurm 等, 2019](#))。

其他与迁移学习相关的研究还包括[Yang 等 \(2016\)](#) 采用迁移学习的方法让分类器学习到不同角度下 SAR 目标的共同特征; [Malmgren-Hansen 等 \(2017\)](#) 利用仿真 SAR 数据预训练神经网络然后迁移到真实的 SAR 数据上。最近的研究开始将域适应方法应用到 SAR 图像的迁移学习中([Geng 等, 2020](#)),在不同域之间迁移时降低特征的差异,增强特征在目标域中的表达能力。

迁移学习作为一种能帮助少量训练样本进行深度学习的有效方法,在 SAR 图像解译中将扮演一个重要的角色。深入探索迁移学习在 SAR 图像解译中的关键性因素能够帮助人工更有效地设计算法。类比于 ImageNet 预训练网络在光学图像界的广泛应用,具有强大的 SAR 图像特征提取能力的预训练模型在 SAR 图像分类任务中也将发挥巨大的作用。

1.2.4 深度学习和物理模型的结合

多数遥感问题都可以用物理模型进行描述,复杂的物理模型一般包含多个参数以及一系列非线性过程,例如在获取地表反射率时需要求解大气辐射校正模型,从物理建模的大气辐射传输方程来看是一个非常复杂的过程, [Zhu 等 \(2018\)](#) 就提出利用深度学习方法从高光谱图像数据中反演地表反射率。SAR 的成像本身就是一个复杂的电磁散射物理过程,对电磁散射过程进行物理建模一直以来困难重重, [Cui 等 \(2019\)](#); [Bao 等 \(2019\)](#) 就采取学习的方式基于大量的 SAR 目标

实测图像利用生成对抗网络模拟新的 SAR 目标图像。数据驱动的深度学习模型和理论驱动的物理模型一般被看做两种不同类型的方法，目前的遥感应用多在大量基准数据条件下用深度学习方法来代替复杂的、难以准确建模的物理模型。[Zhu 等 \(2017\)](#); [Reichstein 等 \(2019\)](#); [Yuan 等 \(2020\)](#) 均在综述性文章中指出在遥感应用中结合物理模型和深度学习方法是未来发展的一个重要趋势，结合的方式包括在一个复杂的物理系统中用深度学习进行仿真来代替其中的一部分难以描述的物理过程，利用深度学习校正物理模型的观测误差，利用深度学习改善物理模型中的参数化，利用物理过程中可靠的规律来限制深度神经网络的学习等。

目前已经有一些研究开展了物理模型和深度神经网络结合的工作，包括用神经网络根据图像感知运动物体的性质 ([Wu 等, 2015](#))；采用条件生成对抗网络从低分辨率的流体流动图像恢复高分辨率数据时，加入流体运动速度和涡量作为输入的物理量为深度神经网络的学习提供正则化约束，并提出物理感知的数据增强方法来避免过拟合 ([Xie 等, 2018](#))。在遥感领域比较有代表性的研究有，[Sun 等 \(2019\)](#) 提出将深度卷积神经网络从 GRACE 卫星数据中学习到的知识融入到全球水文学模型中，以提升水文学模型的预测准确率；[Karpatne 等 \(2017\)](#) 提出物理引导的神经网络模型用于湖水温度建模。

在深度学习中加入物理模型不仅是为了在准确率上获得提升，更是为了让“黑盒子”的深度神经网络具有一定的物理可解释性，如图1.3所示。这也是第1.2.2节提到的 SAR 图像分类深度学习方法第三个阶段的研究方向。

在 SAR 图像分类相关的应用中，目前涉及到这类方法的研究还较少。在目标识别领域，有的传统研究手段采用电磁散射模型来进行 SAR 目标的匹配识别，如[文贡坚 等 \(2017\)](#) 将复杂目标看作不同散射部件的组合，通过电磁散射参数化建模方法进行基于模型的 SAR 目标识别；[周雨 等 \(2015\)](#) 用散射电磁模拟的方法提取待测图像和电磁仿真图像的特征并进行搜索匹配。近年来有一些研究将生成对抗网络应用到 SAR 图像的仿真上代替电磁散射的参数化建模 ([Guo 等, 2017](#); [Wang 等, 2018a](#); [Cui 等, 2019](#); [Bao 等, 2019](#)) 从而实现数据增强，其中[Wang 等 \(2018a\)](#) 的研究与[Xie 等 \(2018\)](#) 在结合深度学习和物理模型的应用上有异曲同工之妙，作者提出 SAR 目标生成对抗网络模型，针对 SAR 目标的成像对方位角非常敏感的特性，将目标的方位角作为先验信息输入生成模型，使网络在生成 SAR 目标时考虑了方位角的影响，获得更稳定的生成结果。

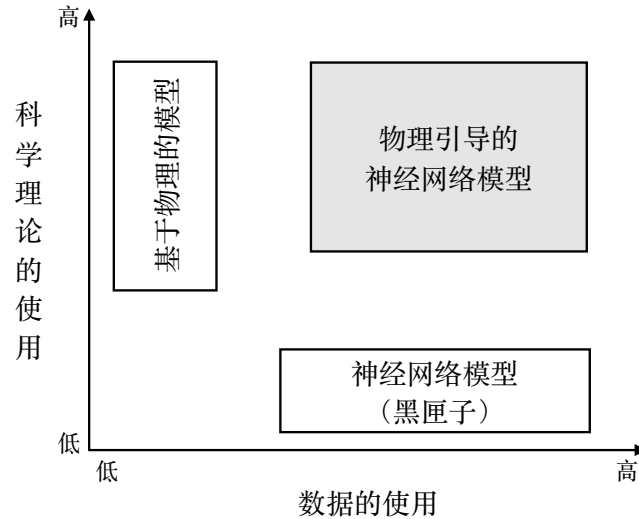


图 1.3 深度学习和物理模型的关系 (Karpatne 等, 2017)

Figure 1.3 The relationship between data-driven deep learning and theory-driven physical models.

在对 SAR 场景的研究中，比较常见的是利用 PolSAR 的目标极化散射矩阵所反映的极化散射模型和极化 SAR 中的统计分布特性逐像素地提取出 SAR 场景的散射特性。在缺乏极化信息的条件下，这种基于物理模型的方法对散射特性的描述会受到限制。Zhao 等 (2019) 提出用复数卷积神经网络直接从全极化 SAR 数据中学习散射特性，并将模型应用到单极化和双极化的数据中；Song 等 (2018) 利用深度神经网络从单极化 SAR 图像中反演全极化 SAR，在视觉上还原全极化 SAR 图像的同时还恢复了图像的极化信息。利用深度学习的方法从单极化 SAR 图像中发掘更多的物理散射特性具有相当的实用性，对 SAR 图像的理解具有重要的意义，除此之外如何在 SAR 场景的分类任务中结合物理散射特性也是今后需要关注的问题。

1.3 研究重点和难点

上述国内外研究现状表明目前在基于深度学习方法的 SAR 图像分类应用中还存在着一些问题亟待解决，在少量样本条件下提高模型泛化能力，根据 SAR 图像与光学图像的不同针对性地开展研究，并结合 SAR 特有的物理散射特性设计深度学习方法，都是未来研究的重点和难点。本文以面向 SAR 图像分类任务的深度学习方法作为研究主题，将拟解决的研究重难点总结为图1.4所示。

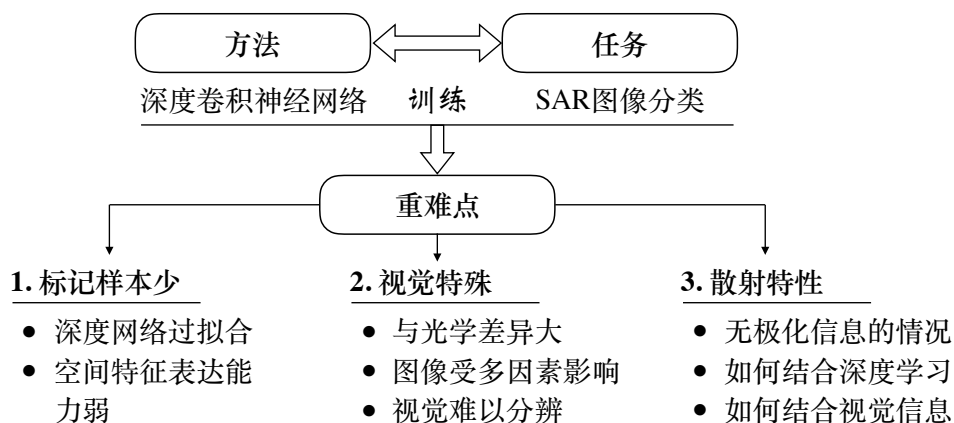


图 1.4 本文的研究重难点分析

Figure 1.4 The research difficulties of the dissertation.

首先，由于 SAR 图像标记较困难，通常情况下都面临训练样本较少的困境，导致训练深度神经网络时出现过拟合，模型的泛化性能较低，对未知 SAR 图像的空间特征表征能力不足。如何在小样本条件下进行有效的空间特征学习，在训练样本减少的情况下能有效地遏制模型性能的下降是当前研究需要关注的重点。针对 SAR 图像分类问题多数情况下都是少量样本的现状，借助一个 SAR 图像空间信息表征能力较强的模型来进行迁移学习，具有重要的实用意义。

其次，SAR 图像在视觉上存在特殊性，和光学图像具有较大差异，例如典型的近距压缩、透视收缩、阴影等表现。常用的基于光学图像的迁移学习手段以及在光学图像的研究中得到的相关结论是否适用于 SAR 图像目前存在一定争议，需要进一步的实验讨论和验证，并提出适合 SAR 图像的迁移策略。另外，SAR 图像呈现的视觉效果与诸多因素有关，包括传感器参数和地面目标的物理性质，因此同种目标或场景在视觉上可能会相差较大，而目视相似的却不一定是同种目标或场景，基于幅度信息的卷积神经网络光凭视觉特性为主可能会产生认知上的困难。

从人类的视觉认知上，SAR 图像反映了地面目标对电磁波的散射强度，除此之外 SAR 图像中同时还包含人眼无法观察的物理散射特性，一般隐含在复数 SAR 图像数据中，例如极化 SAR 的极化散射矩阵。SAR 图像的物理散射特性与地物类别和目标形态有关，因此理解物理散射特性对 SAR 图像分类起着积极的作用。在缺乏极化信息的情况下，基于深度学习的 SAR 图像分类方法很少考虑散射特性，如何借助深度学习方法从复数 SAR 图像中提取非视觉信息，学习物

理散射特性，并应用到 SAR 图像分类任务中，是本文的一大研究重点和难点。

1.4 研究内容与主要贡献

为了提升小样本条件下 SAR 图像空间特征的特征能力，并挖掘 SAR 图像中有用的非视觉信息帮助实现 SAR 图像分类，提升深度模型的稳定性和物理可解释性，本文开展了以下三个部分的研究。

第一部分研究内容：基于迁移学习的 SAR 幅度图像空间特征提取研究，分析迁移学习在 SAR 图像应用中的特殊性，并提出适用于 SAR 图像的迁移学习方法和预训练模型；

第二部分研究内容：单极化复数 SAR 图像的非视觉特性探究，基于时频分析理论提出 SAR 图像的高维时频信号以及目标的子带散射图概念，并设计深度学习算法无监督地学习物理散射特性。

第三部分研究内容：结合 SAR 图像非视觉特性的分类方法研究，包括空间特征和频域特征的融合，以及物理散射主题属性约束下的学习方法。

研究围绕多种 SAR 图像分类问题展开，包括目标识别、地物分类等。表1.1中展示了这三部分的研究内容分别对应的章节，每部分研究内容要解决的问题和解决思路，以及相关的实际应用。

表 1.1 三大研究内容在对应章节、待解决问题、解决思路和相关应用层面上的总结

Table 1.1 The research contents and the corresponding chapters, problems, solutions, and applications.

研究内容	对应章节	待解决问题	解决思路	相关应用
幅度图像的 空间特征学习	第二章	训练样本少	迁移源选择策略	目标识别
	第三章	不同于光学的纹理	传递式、域适应 预训练网络	地物分类
复数图像的 非视觉信息提取	第四章	无极化信息	时频分析理论	物理散射特性提取
	第五章		无监督学习	
面向视觉和 非视觉信息的 分类方法	第四章	训练样本少 视觉上难区分	时频域特征融合	地物分类
	第六章		散射主题属性约束 半监督学习	目标识别 精细化场景分类 海冰分类

第一部分研究包括第二章和第三章的内容，提高深度神经网络对 SAR 图像

空间特征的表征能力。由于在实际应用中,大多数情况下 SAR 图像的标记样本都比较少,仅依靠标记的 SAR 图像很难训练出泛化能力强的模型,如何在小样本条件下进行 SAR 图像的分类是目前受关注较多的一个方向。这部分的研究内容主要关注如何得到在 SAR 图像上泛化能力较强的空间特征表示,探索适用于 SAR 图像的迁移学习方法。第二章提出定量化的特征迁移性分析方法对迁移源的选择进行了讨论,并根据 SAR 图像与光学图像的不同,提出了适用于 SAR 图像的迁移学习方法。第三章则根据第二章的分析结果和迁移学习策略,提出了一个在 SAR 图像上具有较强的空间特征提取能力的预训练模型,并公开发布,为遥感社区 SAR 图像相关方向的研究提供便利。该部分的主要研究成果已分别发表在遥感类期刊 *Remote Sensing*、*IEEE GRSL* 以及顶级期刊 *IEEE TGRS* 上。

第二部分研究包括第四章的第二节和第五章的内容,借助复数 SAR 图像的时频分析理论挖掘潜在的非视觉信息和物理散射特性。多数研究从极化 SAR 图像出发,利用极化目标分解等手段逐目标分析 SAR 图像的物理散射特性,但在缺乏极化信息的情况下,极化分析的方法受到诸多限制。第二部分的研究主要针对单极化的 SAR 图像展开,根据 SAR 特殊的成像机理,SAR 图像在距离向和方位向的频域信号具有不同的物理意义。其中第 4.2 节借助二维连续子带分解得到跨越空间维度和频率维度的 SAR 图像时频信号,这种非视觉信息为复数 SAR 图像的理解提供了不同的观测角度。第五章则进一步提出目标子带散射图的概念,表征了目标在不同方位角和发射频段的散射变化情况,通过无监督学习可以得到子带散射图潜在的特征以及子带散射类别,为理解 SAR 图像的物理散射特性提供了新的视角。该部分的主要研究成果已投稿到遥感类顶级期刊 *IEEE TGRS* 上。

第三部分研究包括第四章的第三节和第六章的内容,将复数 SAR 图像中提取出的非视觉信息与幅度图像中的视觉信息相结合,融合物理散射特性和图像空间特征进行学习。这部分的内容结合了前两部分的研究结果,提出了 SAR 图像分类的深度学习新方法新范式,具有较强的创新性。其中第 4.3 节设计了一个端对端的神经网络模型,将频域信号的特征和空间特征相融合进行协同分类,在区分散射复杂的人造目标以及视觉上相似的场景时具有明显的优势。第六章则直接利用 SAR 图像的物理散射特性作为目标固有的属性来约束空间特征的表达,学习散射主题属性、空间信息和地物语义之间的映射关系。这里的散射特性既

包括极化分析的散射分类结果，也包括单极化的子带散射特性结果，对不同的 SAR 数据具有普适性，并且这种以无监督学习为主，监督学习为辅的方法可以应用到极少标记样本的情形中，为 SAR 图像分类方法提供了一种新的思路。该部分的主要研究成果发表于遥感类顶级期刊 ISPRS JPRS，被国际会议 IGARSS 2020 接收为口头报告。

1.5 论文结构安排

本文共包括七个章节，组织结构如图1.5所示。

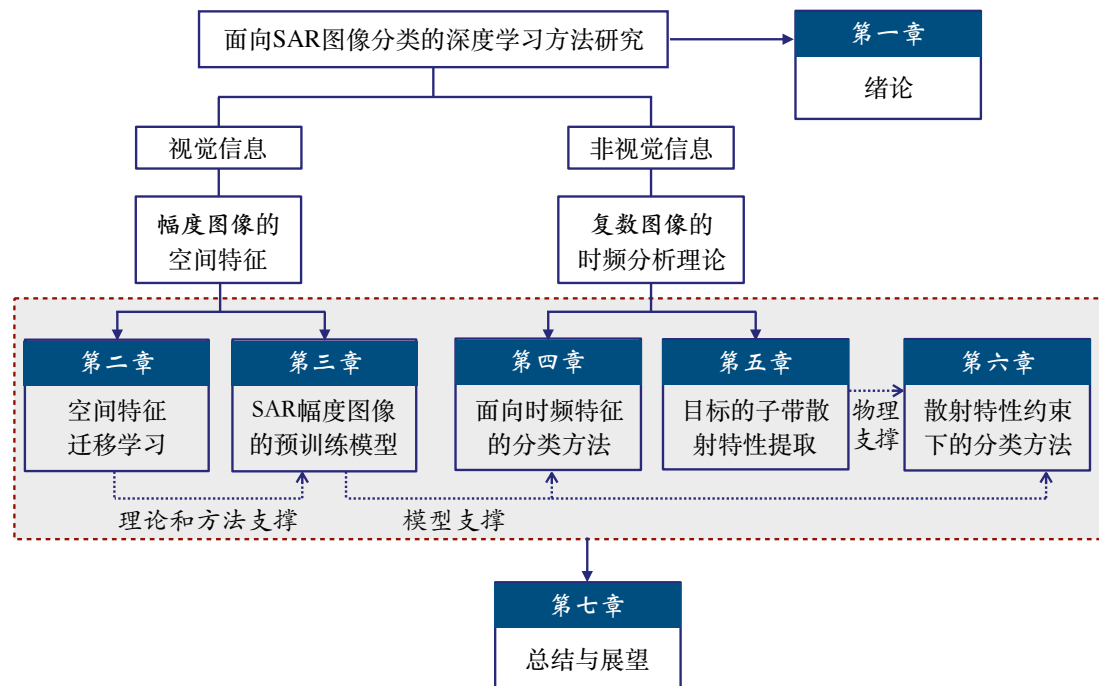


图 1.5 本文的章节组织结构

Figure 1.5 The structure of the dissertation.

第一章为绪论部分，介绍了本文的研究背景及意义，调研了有关“SAR 图像分类的深度学习方法”的国内外研究现状并进行综述，分析了目前的研究重点和难点，并概述了本文的研究内容。

第二章主要研究了 SAR 幅度图像的空间特征迁移学习，包括分析不同迁移源造成特征迁移性的变化，总结出更适合 SAR 图像的迁移学习策略，提出改善不同层次的特征迁移性的方法，为迁移学习在 SAR 图像中的高效应用提供了理论和方法上的支撑。

第三章依托德国宇航中心提供的大规模 SAR 场景标注数据集，训练了一个

深度神经网络预训练模型，能提取丰富的 SAR 图像空间特征，在 SAR 图像上的迁移能力较强，在实际应用中可以作为其他多种 SAR 图像分类任务的预训练网络。

第四章借助二维连续子带分解，将复数 SAR 图像的二维空间信息扩展到高维的时频域信息，包含了 SAR 图像在不同频段的散射表现，并设计深度神经网络自动学习信号的频域特征，辅助空间纹理特征进行协同分类，改善了仅依靠空间信息有时难以区分复杂散射目标的问题。

第五章针对单极化复数 SAR 图像，提出一种基于时频分析理论的子带散射特性无监督深度学习方法，利用 SAR 特殊的成像原理，提取所有目标的子带散射图并学习潜在的特征进行分类，为理解单极化 SAR 图像的物理散射特性提供了一种新的视角，也为单极化 SAR 图像的理解提供了物理层面的支撑。

第六章将已有的散射特性分类结果图作为物理层面的先验信息，初步提出一种散射特性约束下的 SAR 图像分类半监督学习框架，将已知的物理散射特性作为 SAR 图像的固有属性去约束空间特征的学习。该框架可以适用于所有极化方式的 SAR 图像，并在极少标记样本的监督下完成学习。

第七章对本文的研究内容和主要贡献作出总结，针对本文尚且存在的不足，以及目前领域发展的趋势，提出几个未来的工作方向，特别指出将物理引导的深度学习方法作为今后研究的主要研究内容。

第2章 空间特征迁移学习

2.1 引言

SAR 图像中包含的空间纹理信息作为 SAR 图像分类的重要依据受到广泛关注。在传统的基于特征的分类方法中,多采用提取 SAR 幅度图像的空间特征并设计分类器进行学习的方法。在空间特征提取上,传统方法一般采用手动提取特征的方法,例如傅里叶描述子 (Anagnostopoulos, 2009)、基于主成分分析 (Principal Component Analysis, PCA) 的特征 (Mehra 等, 1998)、多尺度的尺寸不变特征变换 (Scale Invariant Feature Transform, SIFT) 算子 (Ruan 等, 2016)、单演信号 (Dong 等, 2014) 等。在分类器选择上,多采用支撑向量机 (Anagnostopoulos, 2009)、贝叶斯分类器 (Mehra 等, 1998)、基于稀疏表示的分类器 (Zhang 等, 2012)、均方根误差分类器 (Novak 等, 2000) 等。除此之外, Yu 等 (2011); Liu 和 Li (2013) 提出相比于单个分类器,多种分类器的融合能进一步提升模型性能。

基于深度神经网络的方法可以直接在数据中通过学习自动得到特征,并采用端到端的模式同时进行特征提取和分类器的学习。但是训练数据的规模是制约模型性能和泛化能力的一个重要因素。相比于光学图像, SAR 目标数据比较匮乏,导致在深度卷积神经网络学习的过程中极易出现过拟合。一些研究采用数据增强 (Morgan, 2015; Ding 等, 2016; Wilmanski 等, 2016) 或设计专门的小型深度卷积网络 (Chen 等, 2016; Lin 等, 2017) 等手段防止过拟合,增强模型在测试数据上的泛化能力。面对训练样本有限的问题,在自然图像或光学遥感图像领域常采用迁移学习的方法,将在大规模数据上训练的模型迁移到小样本任务中,其中 ImageNet 大规模数据集被广泛用作各类光学遥感图像任务的迁移源,例如图像配准 (Wang 等, 2018b)、飞机检测 (Chen 等, 2018)、场景分类 (Marmanis 等, 2016; Zhao 等, 2017a)、遥感图像分割 (Fu 等, 2017) 以及超分辨重建 (Yuan 等, 2017) 等。由于光学遥感图像呈现出地物的光波反射特性,与自然图像在视觉上具有很大的相似性,因此在光学遥感图像的应用中,从自然图像的预训练网络进行迁移的方法取得了一系列的成功。

相比于光学遥感图像,迁移学习在 SAR 图像上的应用目前具有一定争议。Kang 和 He (2016); Wang 等 (2017); Zhong 等 (2019); Wang 等 (2019); Lu 和

Li (2019) 针对 SAR 图像数据, 选择直接在光学图像的预训练网络上进行微调。Marmanis 等 (2017); Wurm 等 (2019); Zhu 等 (2017) 则认为由于 SAR 与光学遥感图像成像机制的不同, 二者在视觉上呈现出很大的差别, 因此基于光学图像训练的模型对 SAR 图像并不具备良好的迁移性, 甚至会削弱深度神经网络对 SAR 图像的特征提取效果。

在 SAR 图像解译领域中, 如何有效地进行迁移学习依旧缺乏足够的讨论。本章以 SAR 目标识别作为焦点任务开展研究, 有针对性地提出适用于 SAR 图像的迁移学习方法, 并对迁移学习中的几个关键性因素进行分析, 提供定量化的实验结果用于制定合适的迁移学习策略, 服务于实际应用和后续研究。

本章第二节提出定量化的特征迁移能力分析方法; 第三节首先提出基于无标签 SAR 图像的迁移学习算法, 然后从任务、数据和网络三个方面对预训练网络的特征在 SAR 目标识别任务中的迁移性进行分析; 第四节提出基于多种迁移源的传递式迁移学习算法, 提升中低层特征的迁移能力; 第五节提出域适应方法, 用于改善高层特征对目标任务的不适应性。

2.2 特征迁移能力分析方法

本节提出定量化的特征迁移能力分析方法, 将深度神经网络各层提取的特征分为一般性 (generic) 和特殊性 (specific) 两类, 用于描述预训练网络在目标任务上的迁移能力。一般性的特征定义为, 其对于目标任务也具有良好的特征表示能力; 特殊性的特征则定义为, 其只对源任务具有良好的特征表示能力, 而对目标任务的特征表示能力不足。特征的一般性强表明特征对目标任务具有良好的迁移能力, 特征具有特殊性则表明特征对目标任务不具有正面的迁移能力。

在 (Yosinski 等, 2014) 的研究基础上, 本节提出定量化特征迁移能力分析方法。假设一共有 n 个不同的源任务, 记为 $S^i, i \in 1, 2, \dots, n$, 目标任务记为 T , 深度神经网络 N 包含 L 个特征层。本章拟分析:

- 源任务 S^i 训练的网络骨架 N 中第 k 层的特征对目标任务 T 的特征迁移能力;
- 对于同样的网络骨架 N , 采用不同源任务 S^i 和 S^j 进行迁移时, 第 k 层特征对目标任务 T 的特征迁移能力;
- 网络 N 各层对目标任务 T 的特征迁移能力随层数 k 的变化。

用源任务 S^i 训练的网络 N 记为 $N(S^i)$ ，用目标任务 T 训练的网络 N 记为 $N(T)$ 。如图 2.1 所示，将 $N(S^i)$ 的前 k 层固定，第 $k+1$ 到第 L 层的参数采用随机初始化，用目标任务 T 来训练，得到网络 $N(S_k^i T)$ 。如果网络 $N(S_k^i T)$ 的表现优于网络 $N(T)$ ，则说明 $N(S^i)$ 的第 k 层特征对目标任务 T 具有一般性，反之如果 $N(S_k^i T)$ 的表现不如 $N(T)$ ，则说明第 k 层特征具有特殊性，对目标任务 T 产生了负向迁移影响 (negative transfer)。

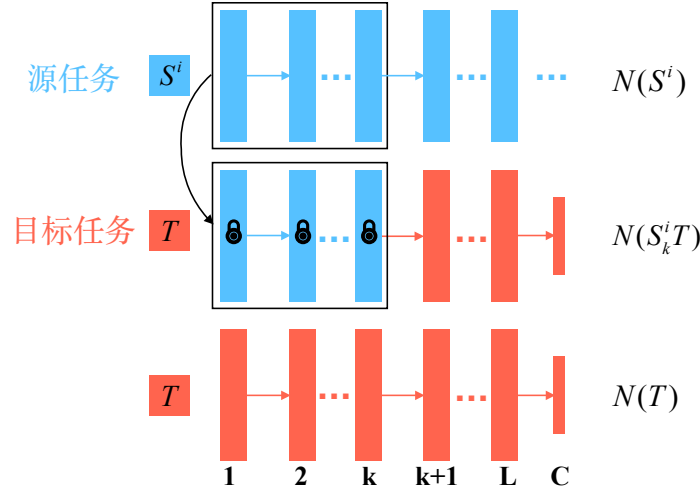


图 2.1 源任务 S^i 第 k 层特征迁移能力分析示意图

Figure 2.1 The analyses of feature transferability in the k th layer of source task S^i .

图 2.2 展示了两个不同源任务在相同网络架构下的特征迁移性分析方法。对于两个不同的源任务 S^i 和 S^j ，迁移并固定前 k 层特征在目标任务 T 上训练的网络分别记为 $N(S_k^i T)$ 和 $N(S_k^j T)$ ，如果 $N(S_k^i T)$ 的表现优于 $N(S_k^j T)$ ，则表明源任务 S^i 的预训练模型第 k 层特征的迁移能力强于源任务 S^j 。

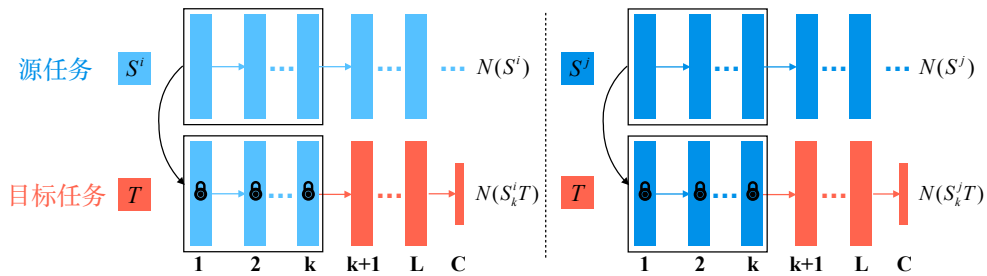


图 2.2 源任务 S^i 和源任务 S^j 的第 k 层特征迁移能力对比分析示意图

Figure 2.2 The analyses of feature transferability in the k th layer of source task S^i and S^j .

在分析过程中依次选择不同的层作为研究对象，由此得到不同预训练网络

$N(S^i)$ 的各层特征对目标任务 T 的迁移能力变化。

2.3 迁移源分析

本章在分析网络各层特征迁移能力时，均采用 OpenSARShip 公开数据集提供的船只样本识别作为目标任务进行实验。OpenSARShip 是Huang 等 (2018) 提出的一个基于 Sentinel-1 卫星数据的船舶数据集，共包含 11345 张船舶切片，来自 41 张 Sentinel-1 在 IW 模式下的卫星图像，极化形式包括 VH 和 VV。数据集共涵盖 17 种类型的船舶，其中在货船 (Cargo) 的分类下还有几种更细的类别。由于数据集中各类别的船只数量分布极其不均衡，某些样本切片过小，为了尽量减少其他因素的影响，并保证切片包含的信息足够清晰有效，本实验从 Cargo 大类下选取三种类别随机挑选均匀的样本数量进行验证，每个样本切片边长均大于 70 像素，如图 2.3和表 2.1所示。

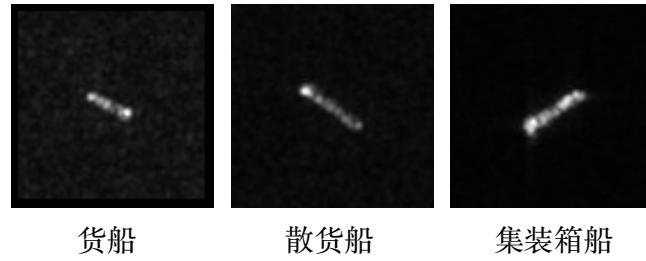


图 2.3 本实验所用 OpenSARShip 数据集的三类目标示意图

Figure 2.3 The examples of three types of ships in OpenSARShip dataset.

接下来，本节将从任务、数据和网络三个方面的差异来分析 SAR 图像分类任务中迁移源的选择策略。

表 2.1 本实验所用的 OpenSARShip 数据集子集

Table 2.1 The applied OpenSARShip dataset in our experiment.

类型	货船	散货船	集装箱船	总计
训练集	100	100	100	300
测试集	79	132	135	346

2.3.1 任务差异

2.3.1.1 无标签 SAR 图像重建任务迁移方法

相比于数量较少的带标记 SAR 目标图像，无标签的 SAR 场景图像数量可观且极易获取。由此，本节提出一种基于无标签 SAR 场景图像的迁移学习方法，将无标签的 SAR 场景图像重建作为源任务迁移到 SAR 目标识别任务中，同时将 SAR 目标图像重建时产生的重建误差引入到分类任务中进行协同优化。在这里设计了一个端对端的深度卷积神经网络结构来实现，集成了重建支路、分类支路和反馈旁路，如图 2.4所示。其中重建支路旨在对输入图像进行尽可能的空间信息恢复，分类支路用于 SAR 目标图像的识别，反馈旁路将重建支路上不同位置得到的重建误差引入分类支路，与分类任务协同作用、联合优化。采用这样的设计是为了让神经网络在重建大量无标签 SAR 图像的过程中学习到丰富的 SAR 图像特征表示，在完成分类任务时将其迁移到 SAR 目标上，同时引入不同位置的重建误差，使得在分类支路最深层仍能保留一部分输入信息，抑制了图像信息在深度网络的流动中出现的梯度消失现象。

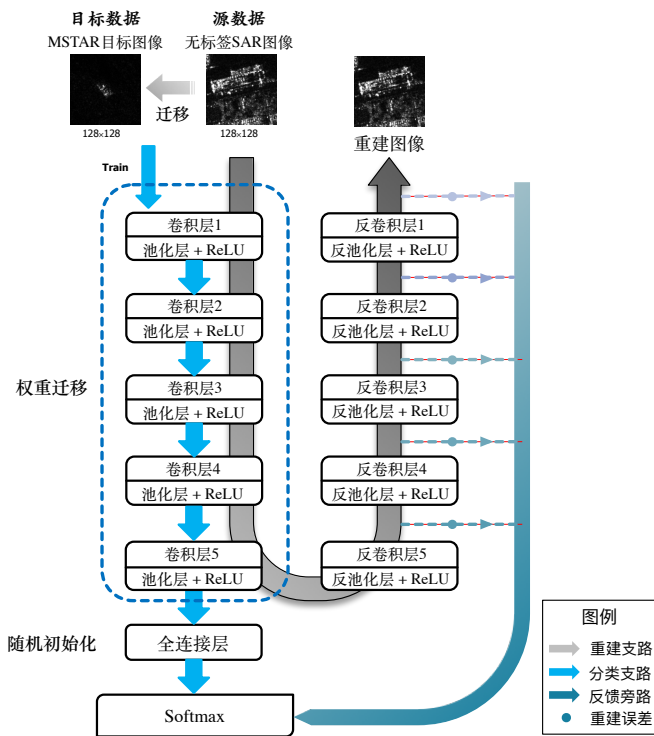


图 2.4 基于无标签 SAR 图像的 SAR 目标识别迁移学习网络架构

Figure 2.4 The transfer learning network for SAR target recognition.

网络的分类支路设计遵循几个通用的标准：采用金字塔形式逐层增加特征

表 2.2 分类支路各层网络设计

Table 2.2 The network design of classification pathway.

层	卷积核大小	通道数	零填充
conv1	5×5	48	2
conv2	5×5	96	2
conv3	3×3	128	1
conv4	3×3	128	1
conv5	3×3	256	1
神经元个数			
fc6		128	
fc7		10	

图的通道数，同时通过下采样减小特征图的大小；采用较小的卷积核，更多地考虑提取图像中的局部信息；采用正则化手段，应用 Dropout 向网络架构中注入噪声来提高泛化能力。分类支路由五个卷积层和两个全连接层组成，如表2.2所示。重建支路由反卷积层和上采样层组成，与分类支路的卷积层一一对应，使得每个反卷积层能将上一个卷积操作生成的特征图还原到卷积之前的状态。反馈回路则从重建回路的每个反卷积层处伸出，将每层的重建误差引入分类损失函数中。

本节采用栈式卷积自编码来完成重建支路的学习，通过堆叠的卷积层对图像进行卷积编码，再将特征图用反卷积层进行重建，由此得到一系列能够提取出丰富特征的卷积核。具体来说，每个卷积层训练的方法较为类似，单元结构如图 2.5所示。

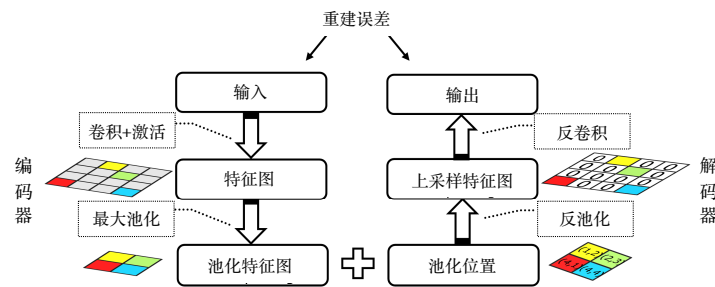


图 2.5 卷积自编码器单元结构示意图

Figure 2.5 The convolutional auto-encoder unit.

对于第 l 层的训练，设输入为带有 M_{l-1} 个通道的特征图 x^{l-1} ，即 x_1^{l-1}, \dots ,

$x_{M_{l-1}}^{l-1}$ ，由第 $l-1$ 层生成。通过编码部分的第 j 个特征图可以表示为：

$$a_j^l = \sum_{i=1}^{M_{l-1}} x_i^{l-1} * K_{ij}^l + b_j^l, j = 1, 2, \dots, M_l \quad \dots (2.1)$$

$$h_j^l = f(a_j^l) \quad \dots (2.2)$$

其中 K_{ij}^l 和 b_j^l 表示第 l 层的第 j 个卷积核和偏置， $*$ 代表二维卷积操作，函数 f 代表整流线性单元 ReLU ， $f(x) = \max(0, x)$ 。为了在卷积操作之后保持特征图大小与输入相同，在卷积之前对输入采用 $\frac{k-1}{2}$ 的 0 填充，这里的 k 表示卷积核的大小。

下采样采用了“what”和“where”的策略，即下采样的结果中不仅保留了采样之后的数据，还保留了采样的位置，便于在解码部分能够更好地恢复图像局部特征。下采样和上采样的数学表达式如下：

$$\text{down}h^l = \mathbb{D}(h^l) \quad \dots (2.3)$$

$$\text{up}h^l = \mathbb{U}(\text{down}h^l) \quad \dots (2.4)$$

上采样之后反卷积层重建的信号可以表示为：

$$y_i^{l-1} = f\left(\sum_{j=1}^{M_l} \text{up}h_j^l * \mathbb{R}(Q_{ji}^{l-1}) + c_i^{l-1}\right), i = 1, 2, \dots, M_{l-1} \quad \dots (2.5)$$

其中 Q 表示反卷积层中的卷积核， $\mathbb{R}(\cdot)$ 意为核 Q 被旋转了 180° ，也就是说矩阵 Q 分别在 x 方向和 y 方向上都做了转置，和卷积层相似，反卷积操作之前也同样使用了 0 填充。

均方根误差在这里被用来衡量重建误差，因此训练第 l 层时所用到的损失函数可以表示为：

$$J(K, b, Q, c) = \text{MSE}^l = \sum_{i=1}^{M_{l-1}} \frac{1}{2N} \|x_i^{l-1} - y_i^{l-1}\|_F^2 \quad \dots (2.6)$$

其中 N 表示第 $l-1$ 个卷积层输出的特征图的大小。 K, Q, b 和 c 这几个参数的值通过最小化损失函数来求得，整个过程通过带有动量和权重衰减的 mini-batch 随机梯度下降方法来实现。损失函数 J 对各参数的梯度由以下公式给出：

$$\frac{\partial J(K, b, Q, c)}{\partial K} = \delta_c^l * \mathbb{R}(x^{l-1}) \quad \dots (2.7)$$

$$\frac{\partial J(K, b, Q, c)}{\partial Q} = \delta_d^l * uph^l \quad \dots (2.8)$$

$$\frac{\partial J(K, b, Q, c)}{\partial c} = \sum_{u,v} (\delta_d^l)_{u,v} \quad \dots (2.9)$$

$$\frac{\partial J(K, b, Q, c)}{\partial b} = \sum_{u,v} (\delta_c^l)_{u,v} \quad \dots (2.10)$$

其中 δ_c^l 和 δ_d^l 分别表示卷积层和反卷积层的误差项。 δ_c^l 和 δ_d^l 可以和采样层的误差项 δ_s^l 联系起来：

$$\delta_s^l = \delta_d^l * Q \odot f'(a^l) \quad \dots (2.11)$$

$$\delta_c^l = \delta_s^l \odot f'(a^l) \quad \dots (2.12)$$

其中 \odot 代表哈达玛积。

参数在迭代过程中根据动量梯度下降原理以 σ 作为学习率进行更新，动量和权重衰减分别设为 0.9 和 0.005。以参数 K 为例：

$$v_{(i+1)} = 0.9 \cdot v_{(i)} - 0.005 \cdot \sigma \cdot K_{(i)} - \sigma \cdot \left\langle \frac{\partial L}{\partial K} \middle|_{K_{(i)}} \right\rangle_{B_i} \quad \dots (2.13)$$

$$K_{(i+1)} = K_{(i)} + v_{(i+1)} \quad \dots (2.14)$$

其中 i 是迭代次数， $\left\langle \frac{\partial L}{\partial K} \middle|_{K_{(i)}} \right\rangle_{B_i}$ 表示第 B_i 批数据的平均偏导数。

本算法采用逐层训练再堆叠的方式来训练整个重建回路，如图 2.6，意味着当前训练的卷积层的输入 x_l 是上一个已经训练好的卷积层的输出 y_{l-1} ，当前训练的反卷积层只需要将信号重建到 x_l 即可。这样一来，每次训练的时候都是一个浅层的网络，相比于更新所有堆叠起来的层的参数更加简单易行。

本方法在完成分类任务时，SAR 目标在重建回路的加权误差被作为正则化项反馈到分类支路中，因此在这里将 SAR 目标识别作为主要的目标任务，SAR 目标重建作为辅助的目标任务。利用无标签 SAR 场景图像得到重建回路预训练网络后，先进行辅助目标任务的迁移。考虑到 SAR 目标的重建和 SAR 场景图像的重建任务在数据和任务类型上都较为类似，因此采取固定分类支路，只微调解码部分的策略。之后再进行主要目标任务的迁移，将全连接层随机初始化，利

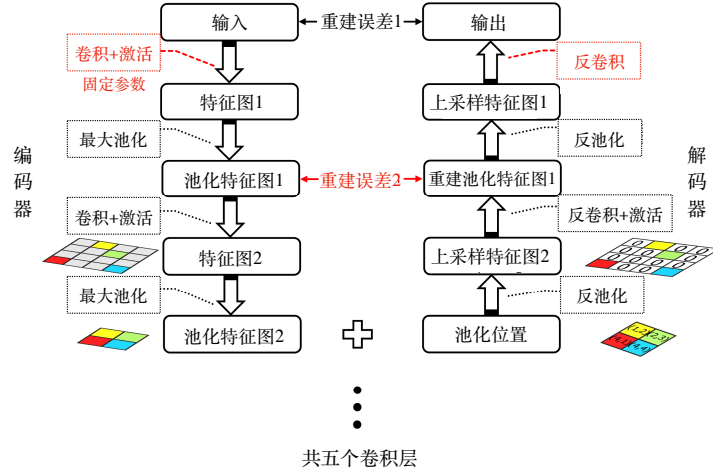


图 2.6 栈式卷积自编码器堆叠示意图

Figure 2.6 The stacked convolutional auto-encoder.

用 SAR 目标数据微调分类支路，同时激活重建回路及反馈回路，总的损失函数可以表示为：

$$Loss = Loss_{recon} + Loss_{cls} \quad \dots (2.15)$$

其中 $Loss_{recon}$ 表示在重建回路中每一层产生的重建误差之和：

$$Loss_{recon} = \sum_{i=1}^5 \lambda_i MSE^i \quad \dots (2.16)$$

而 $Loss_{cls}$ 则表示标签和最后一层的 Softmax 激活值输出之间的交叉熵。 $Loss_{recon}$ 保留了一部分中间层和输入的信息，不同的 λ_i 的值可以控制每一层的误差在最终的误差中所占比例。本方法采用大量的无标签 SAR 图像进行预训练，使网络预先具有一定的 SAR 图像特征提取能力，在迁移到 SAR 目标进行训练的同时引入重建误差，作为一种正则化处理，有助于防止过拟合，提升模型的泛化能力。

2.3.1.2 实验结果

本节对上小节提出的方法进行实验验证。选取 TerraSAR-X 卫星 3 米分辨率的条带模式数据作为无标签的 SAR 图像源，共计三万张切片，并将 MSTAR 识别任务作为目标任务。MSTAR 目标数据集 ([Defense Advanced Research Projects Agency](#)) 是美国国防高等研究计划署 (DARPA) 推出的 SAR 目标识别公开数据集，由工作在 X 波段的聚束模式 SAR 传感器对地面静止的车辆目标进行成像，分辨率为 0.3 米，极化方式为 HH。MSTAR 数据集提供了目标在不同条件下覆盖 0 到 360° 方位角的成像结果，在标准工作条件下 (Standard Operation Condition,

SOC) 俯视角 17° 采集到的目标被作为训练集, 俯视角 15° 的目标作为测试集。训练集和测试集共包括十类军事目标, 分别是自行榴弹炮 2S1、步兵战车 BMP2、装甲侦察车 BRDM2、装甲运输车 BTR60、装甲运输车 BTR70、推土机 D7、坦克 T62、坦克 T72、货运卡车 ZIL131、自行高炮 ZSU234。具体如表 2.3 所示:

表 2.3 MSTAR 数据集

Table 2.3 MSTAR Dataset

类别	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU23	总计
训练集 (17°)	299	233	298	256	233	299	299	232	299	299	2747
测试集 (15°)	274	195	274	195	196	274	273	196	274	274	2425

将上节所提方法记作 CNN-TL-bypass, 表 2.4 展示了 CNN-TL-bypass 与近几年较经典的 SAR 目标识别算法在 MSTAR 数据集上的表现, CNN-TL-bypass 算法和 SVM、MSRC (Dong 等, 2014)、TJSR (Dong 等, 2015)、SDDL SR (Song 等, 2016) 和 JDSR (Sun 等, 2016) 相比分别提升了 8.67%、5.43%、5.68%、2.85%、4.09%, 和基于深度学习的目前最好方法相比 (截止 2017 年 8 月) 同样具有竞争力 (A-ConvNet (Chen 等, 2016) 和 DCHUN (Lin 等, 2017))。

表 2.4 SAR 目标识别算法在 MSTAR 数据集上的测试结果

Table 2.4 The test accuracies of SAR target recognition methods on MSTAR dataset.

方法	SVM	MSRC (Dong 等, 2014)	TJSR (Dong 等, 2015)	SDDL SR (Song 等, 2016)
准确率	90.42%	93.66%	93.41%	96.24%
方法	JDSR (Sun 等, 2016)	A-ConvNet (Chen 等, 2016)	DCHUN (Lin 等, 2017)	CNN-TL-bypass
准确率	95%	99.13%	99.09%	99.09%

将直接用 MSTAR 数据集训练的分类支路记为 CNN-baseline, 只进行迁移学习但未采用反馈旁路的结果记为 CNN-TL。对比图 2.7 中所示的 CNN-baseline 和 CNN-TL 在训练过程中损失函数的变化情况, CNN-baseline 情况下损失函数一开始出现平台期, 难以找到梯度下降的方向, 模型收敛较缓慢, 并在后期出现较为严重的过拟合, 使其在测试集上的泛化能力较弱; 相比之下 CNN-TL 采用迁移学

习使得模型具有较好的初始化参数，损失函数下降较快，过拟合问题有所减轻。

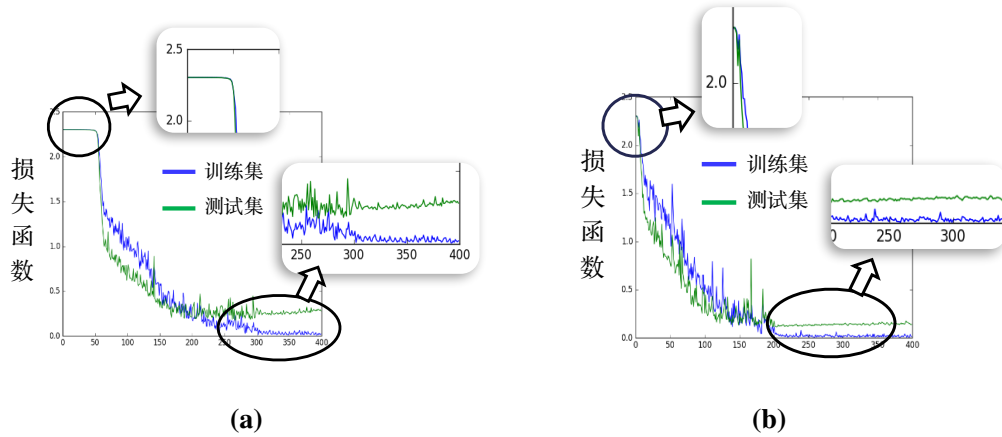


图 2.7 CNN-baseline 和 CNN-TL 训练集和测试集损失函数变化情况。(a) CNN-baseline, (b) CNN-TL.

Figure 2.7 The loss curves of training and test dataset in CNN-baseline and CNN-TL. (a) CNN-baseline, (b) CNN-TL.

CNN-baseline 能达到 98.3% 的识别率，见表 2.5，相比之下基于迁移学习方法的 CNN-TL 能将结果提升到 99.05%，见表 2.6。

表 2.5 CNN-baseline 模型 MSTAR 目标分类结果混淆矩阵

Table 2.5 The confusion matrix of MSTAR target recognition on CNN-baseline model.

类别	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234	准确率
2S1	270	0	0	0	0	0	3	0	0	1	98.54%
BMP2	0	180	0	0	0	0	0	15	0	0	92.3%
BRDM2	0	3	270	0	0	0	0	1	0	0	98.54%
BTR60	0	1	0	195	0	0	0	0	0	0	99.49%
BTR70	1	0	1	0	182	0	2	9	0	0	93.33%
D7	0	0	0	0	0	272	0	0	0	0	100%
T62	0	0	0	0	0	1	272	0	0	0	99.63%
T72	0	0	0	0	0	0	1	195	0	0	99.49%
ZIL131	0	0	0	0	0	0	0	0	274	0	100%
ZSU234	0	0	0	0	0	0	0	0	0	274	100%
准确率	98.3%										

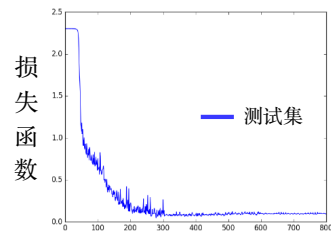
加入反馈旁路后的训练模型记为 CNN-TL-bypass，在 CNN-TL 训练过程中已经收敛的网络，引入加权的各层重建误差之后损失函数又开始下降，如图 2.8 所示，因此 CNN-TL-bypass 在每类 50 个样本用于训练的情况下相比 CNN-TL，识别率能提升 0.49%。图 2.9 展示了在样本量逐渐减少的情况下 CNN-TL-bypass 和

表 2.6 CNN-TL 模型 MSTAR 目标分类结果混淆矩阵

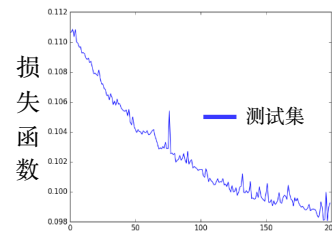
Table 2.6 The confusion matrix of MSTAR target recognition on CNN-TL model.

类别	2S1	BMP2	BRDM2	BTR60	BTR70	D7	T62	T72	ZIL131	ZSU234	准确率
2S1	274	0	0	0	0	0	0	0	0	0	100%
BMP2	0	194	1	0	0	0	0	15	0	0	99.49%
BRDM2	1	2	271	0	0	0	0	0	0	0	98.9%
BTR60	0	0	0	196	0	0	0	0	0	0	100%
BTR70	1	1	0	5	182	0	4	1	1	0	93.3%
D7	0	0	0	0	0	271	1	0	2	0	98.9%
T62	0	0	0	0	0	1	272	0	0	0	99.63%
T72	0	1	0	0	1	0	0	194	0	0	98.9%
ZIL131	0	0	0	0	0	0	0	0	274	0	100%
ZSU234	0	0	0	0	0	0	0	0	0	274	100%
准确率	99.05%										

其他方法相比的结果，验证了 CNN-TL-bypass 方法在少量训练数据的情况下能保持较为稳定的性能。



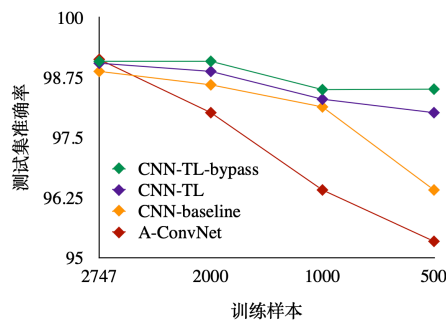
CNN-TL测试集损失函数变化



收敛后的CNN-TL加入反馈旁路后测试集损失函数变化

图 2.8 CNN-TL 模型加入反馈旁路前后测试集损失函数的变化情况

Figure 2.8 The test loss curve of CNN-TL model training before and after attaching the bypass.



训练样本数	2747	2000	1000	500
CNN-TL-bypass	99.09	99.09	98.5	98.51
CNN-TL	99.05	98.88	98.3	98.02
CNN-baseline	98.88	98.6	98.14	96.41
A-ConvNet	99.13	98.02	96.41	95.34

图 2.9 训练样本减少情况下 MSTAR 目标识别结果 (%)

Figure 2.9 The performances of MSTAR target recognition with reducing training data (%).

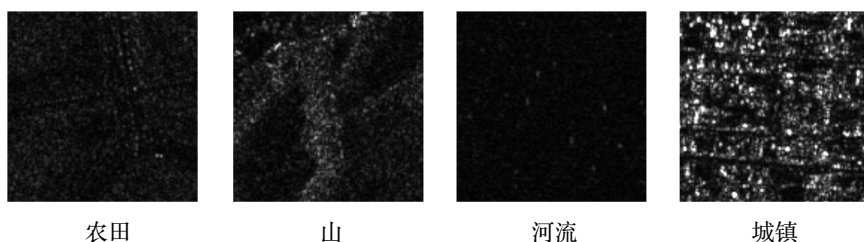


图 2.10 无标签 SAR 场景图像示例

Figure 2.10 The examples of the unlabeled SAR scene images.

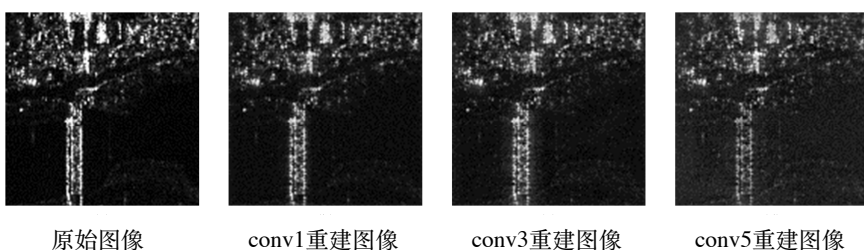


图 2.11 SAR 场景图像在堆栈自编码器各层训练过程中的重建结果

Figure 2.11 The reconstruction results for each layer in SCAE training processing.

2.3.1.3 不同任务的迁移

上一小节的实验结果验证了 SAR 图像重建任务迁移到 SAR 目标识别任务的可行性，且在逐步减少训练数据的情况下能取得良好的效果。第2.3.1.1小节所提的迁移学习算法中包括从 SAR 场景图像重建任务到 SAR 目标重建任务的迁移，以及从 SAR 场景图像重建任务到 SAR 目标分类任务的迁移，本节首先对这两种情况进行对比讨论。

图 2.10展示了无标签 SAR 场景图像数据集中的几个样本示例，包括山脉、城镇、农田、河流等，以图 2.11的示例图片为例，经过逐层训练的堆栈卷积自编码模型各层对输入 SAR 图像的重建效果如图 2.11的 (b)(c)(d) 所示，可以看出堆栈卷积自编码模型在源任务上获得了较为充分的训练。

实验在进行 MSTAR 数据集重建任务的迁移时，首先尝试了在 SAR 场景图像训练的重建回路模型基础上微调所有的层，在这个过程中，卷积核 L2 范数的平均变化值被用来衡量每个卷积层的变化情况。从表 2.7中可以发现，从 SAR 场景重建到 SAR 目标重建的迁移过程中，每层卷积核的变化程度都较小，在百分之零点几左右。变化最大的第一个卷积层可视化结果展示在图 2.12的左侧所示，视觉上能观察到重建任务迁移时卷积核的变化并不明显。因此在 MSTAR 目标图

像重建任务的实际应用中，可以直接使用预训练网络中卷积层的特征表示来提取有效的编码特征，只微调解码部分的反卷积层。表 2.8则展示了 MSTAR 数据集分类任务迁移过程中每层卷积核的 L2 范数变化情况，实验发现第一、二、五层都出现了 L2 范数变化超过 30% 的卷积核，其中第一层卷积核的变化最为明显，迁移前后的第一层卷积核示例如图 2.12右侧所示，与左侧相对比可以观察到在分类任务迁移时卷积核变化更明显，因此在目标分类任务迁移的实际应用中没有直接使用预训练网络的特征，而是采取了卷积层微调的手段。

表 2.7 目标重建任务迁移过程中每层卷积核 L2 范数的变化百分比

Table 2.7 The average kernel variation(%) of transferring reconstruction task in L2-norm of each convolutional layer.

层 (卷积核个数)	conv1(48)	conv2(96)	conv3(128)	conv4(128)	conv5(256)
卷积核 L2 范数的变化 (%)	0.36	0.07	0.1	0.1	0.1

表 2.8 目标分类任务迁移过程中每层卷积核的 L2 范数变化情况

Table 2.8 The kernel variations of transferring classification task in L2-norm of each convolutional layer.

层 (卷积核个数)	conv1(48)	conv2(96)	conv3(128)	conv4(128)	conv5(256)
卷积核 L2 范数变化超过 30% 的个数	37/48	5/96	0/128	0/128	7/256

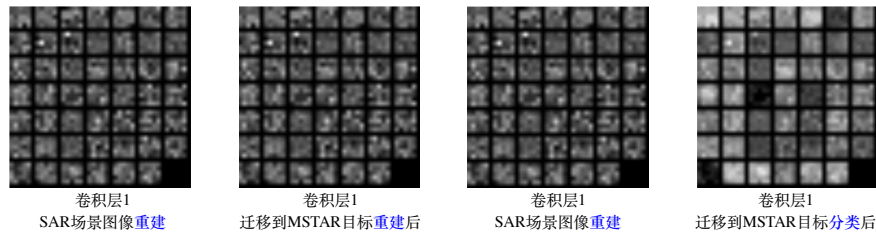


图 2.12 重建任务和分类任务迁移时第一层卷积核的变化情况

Figure 2.12 The convolutional kernels in the first layer before and after transferring reconstruction and classification tasks, respectively.

上面的分析表明与源任务相似的目标任务在进行迁移时，预训练网络的卷积核能提供较好的特征表达能力，因而在缺乏 SAR 图像数据使得微调网络都存在困难的情况下，可以通过直接复用任务类型一致且源数据相似的预训练网络

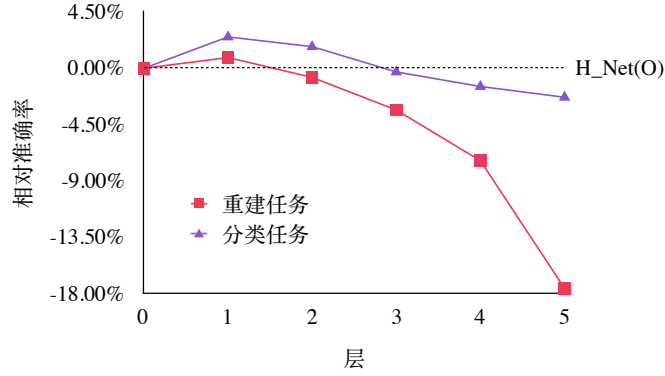


图 2.13 不同源任务迁移到 OpenSARShip 目标识别任务上时网络各层特征的迁移性变化

Figure 2.13 The performances of transferring different layers of reconstruction and classification task to OpenSARShip target recognition task.

中的层进行特征提取，反之则需要更多的 SAR 图像数据对网络各层的参数进行调整。

基于 SAR 船舶目标识别作为目标任务的研究对象，本节还分析了任务差异对预训练网络各层特征迁移性的影响。利用 OpenSARShip 数据集直接训练上节所提网络框架的分类支路，记为 $H_Net(O)$ 作为基准模型，按第2.2小节所提的分析方法依次冻结重建任务和分类任务预训练模型的各层特征进行迁移学习，得到 OpenSARShip 船舶目标相对于 $H_Net(O)$ 的识别准确率，记录于图 2.13中。虚线上方的点表示对应层特征具有较好的一般性，能有效迁移到 SAR 舰船目标识别任务中，虚线下方的点则表示对应层特征对于源任务具有特殊性，不能有效地迁移到 OpenSARShip 目标上。通过图 2.13可以观察到从 SAR 图像重建任务到 SAR 目标识别的迁移过程中，由于任务类型相差较大，除底层特征具有较好的迁移能力之外，其他层特征随着层数加深其对于源任务的特殊性逐步增强，在目标识别任务中无法产生正向的迁移效果。

通过本节的分析可以得出，在 SAR 图像上进行特征的迁移学习时，若参与迁移的源任务与目标任务类型相差较大时，特征对目标域 SAR 图像的表征能力会受到很大的限制，直接复用已有的特征不能取得理想的迁移效果，因此卷积核需要较大程度的调整。但由于重建任务和分类任务均针对 SAR 图像数据，在缺乏大规模标记 SAR 图像的情况下，以此作为初始模型也能为训练提供较好的初始化条件，有利于降低训练成本，与重新设计并从头训练网络骨架相比，此迁移模型对小样本的鲁棒性更强。

2.3.2 数据差异

本节选择相同的网络架构和任务类型，以 OpenSARShip 船舶目标识别任务为目标讨论不同迁移源数据对各层特征迁移性的影响。实验采用 AlexNet 网络的卷积层架构和图像分类任务，选取了三种不同的迁移数据源：



图 2.14 三种迁移数据源

Figure 2.14 The different source data.

1) ImageNet 光学数据集，记为 I

ImageNet 是由Deng 等 (2009) 提出的大规模自然图像目标数据集，计算机视觉领域中应用广泛，一般采用包含 1000 类、约 120 万张图片的子集作为基准，进行目标识别、检测等研究。经过 ImageNet 数据集训练得到的网络，例如 AlexNet(Krizhevsky 等, 2012)、GoogLeNet(Szegedy 等, 2015) 和 ResNet(He 等, 2016) 等，被广泛应用于其他光学图像视觉任务中作为预训练模型使用。

2) TerraSAR-X 场景数据集，记为 S

Dumitru 等 (2016) 提出的 TerraSAR-X 场景数据集包含上万张带标记的 SAR 场景图像切片，具体描述见第3.2节。

3) MSTAR 目标数据集，记为 M

MSTAR 目标数据集包含十类车辆/军事目标，与 OpenSARShip 数据集中的样本类似，每个切片中均只包含一个 SAR 目标。具体描述参考表2.3。

本节选取 AlexNet 网络架构 (Krizhevsky 等, 2012)，考虑到全连接层的参数过多，难以应用到样本较少的 SAR 图像中，因此本节只采用 AlexNet 的卷积层作为主干网络 (backbone)，记为 AlexNet_Conv. 直接用 OpenSARShip 数据集训练 AlexNet_Conv 网络得到的模型结果记为 $AlexNet_Conv(O)$ ，AlexNet_Conv 在 ImageNet 数据集、SAR 场景数据集、和 MSTAR 数据集上的预训练模型分别记为 $AlexNet_Conv(I)$ 、 $AlexNet_Conv(S)$ 和 $AlexNet_Conv(M)$ 。图 2.15展示了采用

不同迁移数据源时,固定不同层 k 进行迁移的实验结果,其中纵坐标表示迁移网络相比于基准模型 $AlexNet_Conv(O)$ 的相对准确率。 $AlexNet_Conv(I)$ 第一层的特征在 OpenSARShip 目标上表现出良好的迁移性,但在冻结第二层以及更高层的特征进行迁移时,基于 ImageNet 的预训练模型的特征对 SAR 目标的迁移性下降十分明显,甚至在第五个卷积层出现了负迁移。相比之下, $AlexNet_Conv(S)$ 和 $AlexNet_Conv(M)$ 各层的特征在 OpenSARShip 数据集上显示出更稳定的一般性。

直观地说,三种迁移源数据与目标数据 OpenSARShip 的相似度排序为 $MSTAR > SAR > ImageNet$,数据集的规模和所含目标的丰富程度排序为 $ImageNet > SAR > MSTAR$ 。ImageNet 数据集虽然在光学自然图像领域展现出强大的迁移能力,广泛应用于各类视觉任务,但由于自然图像和 SAR 图像之间的较大差异,ImageNet 的预训练网络在中高层的特征迁移到 SAR 图像上时会受到限制,表现出不适应性,但底层特征依旧具有比较好的迁移性。尽管另外两种迁移数据源的规模都远不如 ImageNet,但由于 MSTAR 目标与 OpenSARShip 数据集最为接近,作为迁移源能提供最相似的特征表示,用现成的 (off-the-shelf) 特征即可在 OpenSARShip 数据集上表现出良好的识别性能,例如 $AlexNet_Conv(M_2O)$ 识别准确率超过 $AlexNet_Conv(O)$ 6.36%。而 SAR 场景图像包含多样性更丰富的样本,能学习到更全面的 SAR 图像特征,各层特征的迁移能力表现得更加稳定。但由于 SAR 图像标记困难,大规模 SAR 图像场景数据集十分有限,难以获取。

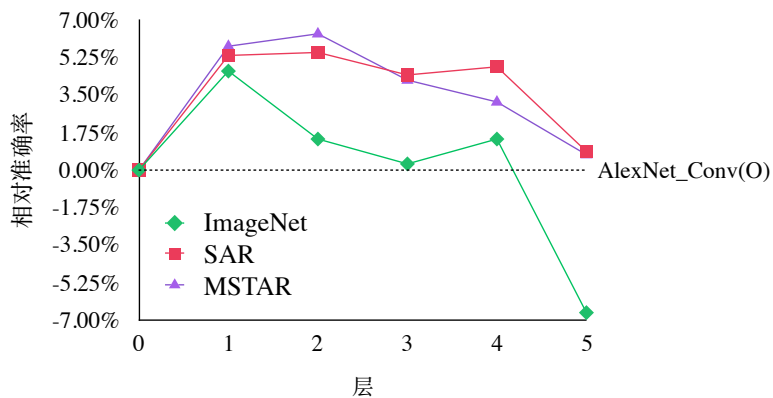


图 2.15 不同迁移数据源在 AlexNet_Conv 网络上预训练后每层特征迁移到 OpenSARShip 目标识别上的表现

Figure 2.15 The performance of transferring different layers in AlexNet_Conv pre-trained model to OpenSARShip target recognition, using different source data.

2.3.3 网络差异

本节讨论不同的深度卷积神经网络在相同的任务类型和迁移数据源条件下会如何影响特征在 SAR 目标识别任务中的迁移性。实验选取最相近的 MSTAR 数据集作为迁移数据源，分类作为迁移任务类型，采用三种网络进行分析讨论，分别是：

- 1) A_ConvNet(Chen 等, 2016), 在 MSTAR 目标识别任务上取得了 99.13% 的准确率；
- 2) H_Net(Huang 等, 2017b), 第2.3.1.1节提出的分类支路结构，在 MSTAR 任务上从头训练能取得 98.3% 的准确率；
- 3) AlexNet_Conv, AlexNet(Krizhevsky 等, 2012) 的前五个卷积层，在 MSTAR 任务上从头训练能取得 97.53% 的准确率。

上述三种网络的卷积层具体结构和参数如表2.9所示， $a(b)$ 表示这一层的卷积核一共有 a 个通道，每个核的大小为 $b \times b$ 。表2.9的最后两行展示了 MSTAR 和 OpenSARShip 数据集在三个网络上从头训练得到的识别率，可以看到随着网络结构的加宽和加深，SAR 目标识别率下降。A_ConvNet(Chen 等, 2016) 是根据 MSTAR 数据集的规模专门设计的一个小型网络，能很好地拟合 MSTAR 数据集中的样本，因此从头训练能获得最优的结果。OpenSARShip 数据集的规模小于 MSTAR 数据集，只包括几百个样本，更宽更深的网络势必会产生过拟合现象。因此相比于更复杂的 H_Net 和 AlexNet_Conv 网络，A_ConvNet 在从头训练 MSTAR 和 OpenSARShip 的情况中都表现最好。

接下来对上述几个预训练模型 $A_ConvNet(M)$ 、 $H_Net(M)$ 和 $AlexNet_Conv(M)$ 的每层特征进行迁移能力分析，如图 2.16所示。尽管 A_ConvNet 能拟合小规模 SAR 目标数据集并取得较优的识别率，但相比于其他两个网络， $A_ConvNet(M)$ 每层的特征对 OpenSARShip 数据集都表现出特殊性，产生了负向迁移影响。而在迁移源 MSTAR 数据集上过拟合最严重的 $AlexNet_Conv(M)$ 的各层特征对 OpenSARShip 目标表现出最优的迁移能力。Azizpour 等 (2015) 提出网络的过参数化，即增加网络的宽度和深度，在迁移到相似的任务上时会提升网络的迁移性。在本节讨论的例子中，虽然更复杂的网络结构会导致在训练源任务 MSTAR 分类时出现过拟合，但迁移前 k 层特征时能表现出更强的一般性。

表 2.9 三种深度卷积神经网络的结构设计及在 MSTAR 和 OpenSARShip 数据集上从头训练的识别率

Table 2.9 The convolutional layer design of three different architectures, and the performances on MSTAR and OpenSARShip dataset.

网络结构	A_ConvNet	H_Net	AlexNet_Conv
conv1	16(5)	48(5)	96(11)
conv2	32(5)	96(5)	256(5)
conv3	64(5)	128(3)	384(3)
conv4	128(6)	128(3)	384(3)
conv5	None	256(3)	256(3)
大小	0.4 M	0.7 M	4 M
MSTAR(%)	99.13	98.3	97.53
OpenSARShip(%)	87.57	85.55	84.39

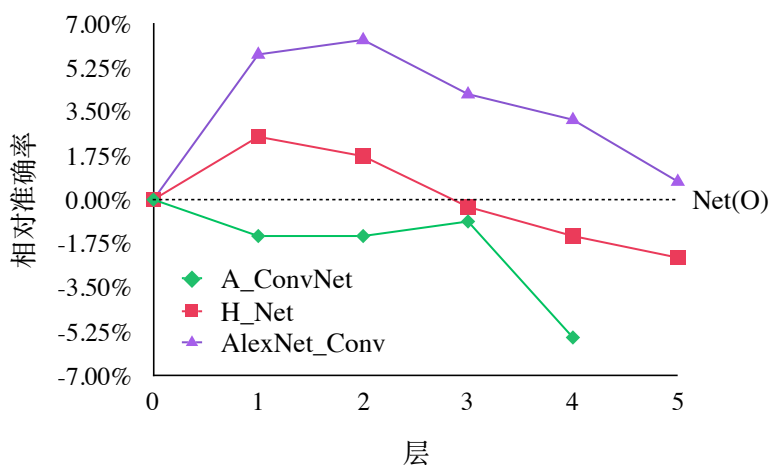


图 2.16 MSTAR 迁移源数据在不同网络上进行预训练后每层特征迁移到 OpenSARShip 目标识别上的表现

Figure 2.16 The performance of transferring different layers in different MSTAR pre-trained models to OpenSARShip target recognition.

2.4 传递式迁移学习

上节对任务、数据和网络三个因素进行了分析, 表2.10总结了针对 OpenSARShip 目标识别任务四种不同迁移源的优势和劣势。由此提出 SAR 图像分类任务中理想化的迁移学习策略为: 借助大规模的带标签 SAR 图像在深层的卷积神经网络上训练分类任务得到预训练模型进行迁移。大规模自然图像数据集与 SAR

目标图像存在的巨大差异，以及大规模的无标签 SAR 场景图像的重建存在任务上的差异，导致了预训练网络的迁移性能受到限制，中高层特征对 SAR 目标分类任务的一般性较差，甚至表现出负迁移现象。小规模的目标图像虽在数据和任务相似性上具有优势，但训练复杂的神经网络较为困难。针对这种现状，本节提出传递式迁移学习方法，利用中间任务来减轻迁移源任务的劣势。

表 2.10 迁移源分析总结

Table 2.10 The summary of transfer sources.

迁移源	优势	劣势
ImageNet 自然图像 (分类)	现成的预训练模型	
	强大的特征提取能力 (光学) 任务相似	数据差异
无标签 SAR 场景图像 (重建)	数据易获取 数据多样性丰富	任务差异
有标签 SAR 场景图像 (分类)	数据相似	
	数据多样性丰富 任务相似	数据难获取
MSTAR 目标图像 (分类)	数据相似	数据少
	任务相似	数据多样性弱

2.4.1 方法介绍

如图 2.17所示，设有目标任务 T 和源任务 S^i ，其中源任务 S^i 包含大规模的数据，但源数据和目标数据相差较大（如 ImageNet）或任务类型不一致（如无标签的 SAR 图像），经源任务训练的网络 N 记为 $N(S^i)$ 。此外，选取其他源任务 S^j, \dots, S^n ，这些任务所包含的数据规模较小，但能弥补源任务 S^i 数据或任务差距较大的缺陷，例如有限的几类 SAR 场景图像，或 MSTAR 目标识别任务。用任务 S^j 在 $N(S^i)$ 上对所有层进行微调以适用于 S^j ，得到网络 $N(S^i * S^j)$ ，以此类推得到模型 $N(S^i * S^j \dots * S^n)$ 。

2.4.2 实验分析

将 ImageNet 数据集（分类任务）和无标签 SAR 图像（重建任务）作为迁移源， $AlexNet_Conv(I)$ 和 $H_Net(S_r)$ 分别表示 ImageNet (I) 和 SAR (S_r) 在 AlexNet_Conv 和 H_Net 上训练的模型。由于二者分别在数据和任务类型上的较

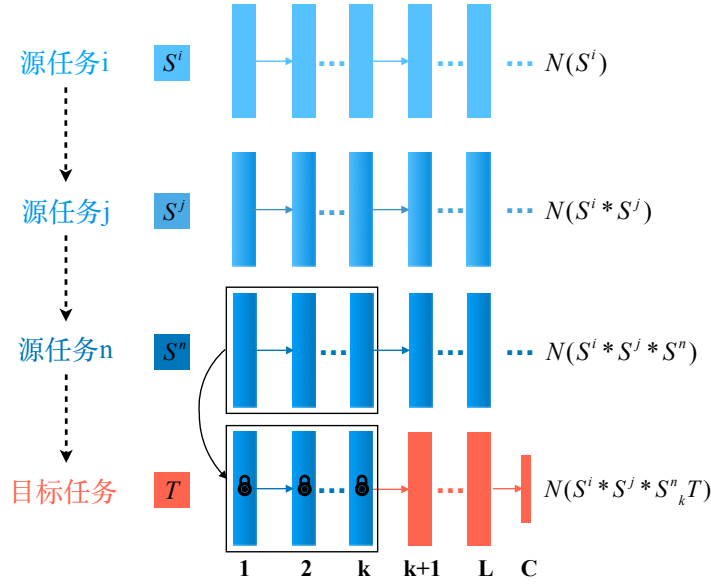


图 2.17 传递式迁移学习示意图

Figure 2.17 The description of transitive transfer learning.

大差异造成除底层之外的特征在 SAR 目标上迁移能力的大幅度下滑，因此本节分别以 $AlexNet_Conv(I)$ 和 $H_Net(S_r)$ 作为初始迁移模型，探究传递式迁移学习算法对特征迁移性的改善效果。

对于 $H_Net(S_r)$ ，将 MSTAR 目标识别作为中间源任务，得到传递式迁移模型 $H_Net(S_r * M)$ ，即第2.3.1.2节中得到的 CNN-TL-bypass 模型。图 2.18展示了 H_Net 网络在传递式迁移学习之后各层特征迁移到 OpenSARShip 识别任务中的表现， $H_Net(S_r * M)$ 明显地改善了 $H_Net(S_r)$ 和 $H_Net(M)$ 在各层，特别是中层和高层特征上的一般性。这说明尽管无标签的 SAR 图像重建和 OpenSARShip 船舶目标识别相比具有任务类型上的差距，但在中间任务 MSTAR 目标识别的帮助下提升了原始模型在目标任务中的可迁移性。

对于 $AlexNet_Conv(I)$ ，考虑 SAR 场景分类和 MSTAR 目标识别两种中间源任务，得到以下几个传递式迁移模型 $AlexNet_Conv(I * S)$ 、 $AlexNet_Conv(I * M)$ 和 $AlexNet_Conv(I * S * M)$ 。图 2.19中的结果表明第二层到第四层的特征在少量 SAR 场景图像 ($AlexNet_Conv(I * S)$) 或 MSTAR 目标数据集 ($AlexNet_Conv(I * M)$) 的帮助下一般性得到了非常明显的提升， $AlexNet_Conv(I * S * M)$ 更是让前四层的特征都维持在一个相当的水平上。还有一个值得关注的现象，在自然图像到 OpenSARShip 的直接迁移过程中，最高层特征表现出极强的特殊性，出现

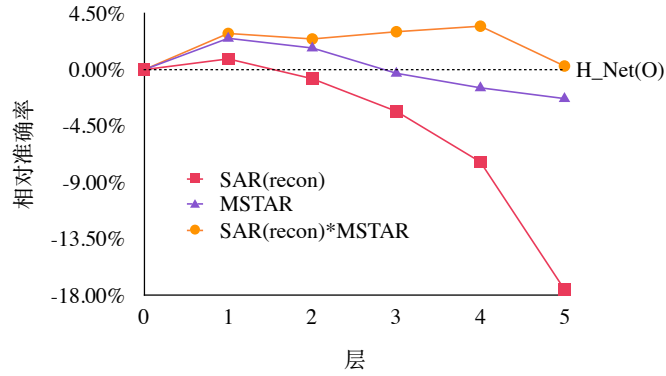


图 2.18 传递式迁移学习前后 H_Net 的每层特征迁移到 OpenSARShip 目标识别上的表现

Figure 2.18 The performance of transferring different layers of H_Net to OpenSARShip target recognition before and after transitive transfer learning.

负向迁移的效果。在中间层特征都随着传递式迁移逐渐提升其在 SAR 目标上的特征迁移能力时，最高层特征的可迁移性在 SAR 目标上变得更差。这个现象说明最原始的迁移源与目标任务是否相差太大会对高层特征的特殊性产生持续的影响，而 SAR 图像和光学图像之间的巨大差异在迁移学习过程中给高层特征带来的负面影响，值得引起相关研究和应用的注意。

上述迁移能力分析所有实验的数值结果记录在表2.11中。

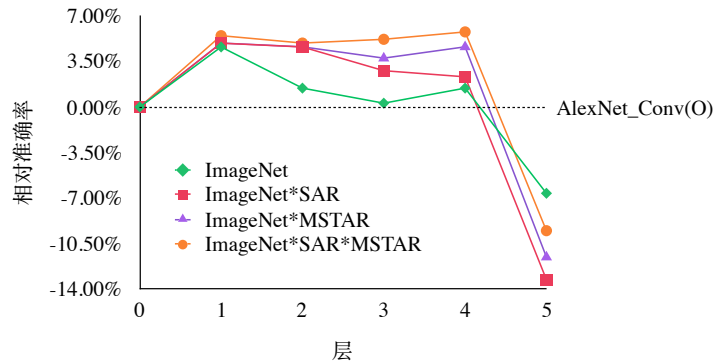


图 2.19 传递式迁移学习前后 AlexNet_Conv 的每层特征迁移到 OpenSARShip 目标识别上的表现

Figure 2.19 The performance of transferring different layers of AlexNet_Conv to OpenSARShip target recognition before and after transitive transfer learning.

表 2.11 采用不同网络和源数据冻结各层特征进行 OpenSARShip 船舶目标识别的迁移时的识别率

Table 2.11 The recognition rate of transferring different layers of different pre-trained models to OpenSARShip target recognition task.

网络	源任务	Net(OpenSAR)	迁移并冻结的层				
			1	2	3	4	5
A_ConvNet	MSTAR	0.8757	0.8612	0.8612	0.8670	0.8208	none
H_Net	SAR(recon)	0.8555	0.8641	0.8483	0.8223	0.7818	0.68
	MSTAR		0.8805	0.8728	0.8526	0.8410	0.8324
	SAR(recon)*MSTAR		0.8844	0.88	0.8858	0.8902	0.8584
AlexNet_Conv	ImageNet	0.8439	0.8901	0.8584	0.8468	0.8584	0.7774
	SAR		0.8974	0.8988	0.8883	0.8921	0.8526
	MSTAR		0.9017	0.9075	0.8859	0.8757	0.8511
AlexNet_Conv (transitive transfer)	ImageNet*SAR	0.8439	0.8930	0.8901	0.8718	0.8671	0.7109
	ImageNet*MSTAR		0.8931	0.8902	0.8815	0.8902	0.7283
	ImageNet*SAR*MSTAR		0.8988	0.8931	0.8960	0.9017	0.7486
	SAR*MSTAR		0.8988	0.9032	0.8959	0.8872	0.8612

2.5 高层特征域适应

前面分析了采用不同迁移源或多种迁移源时网络各层特征的迁移能力,图 2.18 和图 2.19 中均显示出特征迁移能力在最高层有明显的下降。高层特征较弱的迁移性主要由源域和目标域的差异导致,网络在训练过程中越高层的特征越接近用于训练的具体任务,源任务的数据集和目标任务的数据集通常具有不同的概率分布,当源域和目标域差异较大时其概率分布的差异也更大,在这种情况下源数据的高层特征空间和目标数据的匹配程度较低,不能有效地进行迁移。针对这一特点,本节提出基于域适应的迁移学习框架,在传递式迁移学习的基础上利用域适应的方法来拉近源数据和目标数据间的距离,改善高层特征在目标域的不适应性。

2.5.1 方法介绍

最大平均差异 (Maximum Mean Discrepancy, MMD) 由 [Borgwardt 等 \(2006\)](#) 首次提出,用于衡量两种数据分布之间的差异。在迁移学习领域中,大多数的域适应方法都用减小 MMD 的思路来缩小源域与目标域之间的差异。假设源域数据 $\{x^s\}$ 的分布是 p , 目标域数据 $\{x^t\}$ 的分布是 q , 对于两个服从不同分布 p 和

q 的数据集 $\mathcal{D}_s = \{(x_i^s, y_i^s)\}_{i=1}^m$ 和 $\mathcal{D}_t = \{(x_j^t, y_j^t)\}_{j=1}^n$ 来说，它们之间的最大平均差异 MMD 定义为

$$MMD(x^s, x^t) = \sup_{\|\phi\|_{\mathcal{H}} \leq 1} [E_{x^s \sim p}[\phi(x^s)] - E_{x^t \sim q}[\phi(x^t)]]_{\mathcal{H}} \quad \dots (2.17)$$

其中 $\phi(\cdot)$ 表示再生希尔伯特空间 (Reproducing Kernel Hilbert Space, RKHS) \mathcal{H} 的单位球函数集合中的元素， $E_{x^s \sim p}[\cdot]$ 表示服从分布 p 的数据集 x^s 的期望。在再生希尔伯特空间中，期望可以被看做一个嵌入 p ，写作 μ_p

$$\mu_p = E_{x^s \sim p}[\phi(x^s)] \quad \dots (2.18)$$

因此，最大平均差异可以看做在 RKHS 中两个概率分布嵌入之间的距离，将其作为衡量源域数据和目标域数据之间差异的指标。MMD 的平方可写成

$$\begin{aligned} MMD^2(x^s, x^t) &= E_{x^s \sim p} \langle \phi(x^s), \phi(x'^s) \rangle_{\mathcal{H}} \\ &\quad + E_{x^t \sim q} \langle \phi(x^t), \phi(x''^t) \rangle_{\mathcal{H}} \\ &\quad - 2E_{x^s \sim p, x^t \sim q} \langle \phi(x^s), \phi(x^t) \rangle_{\mathcal{H}} \end{aligned} \quad \dots (2.19)$$

其中 $\langle \cdot \rangle_{\mathcal{H}}$ 表示 \mathcal{H} 中的内积运算，特征图 $\phi(\cdot)$ 可以和 \mathcal{H} 中的核映射联系起来，写作 $k(x^s, x^t) = \langle \phi(x^s), \phi(x^t) \rangle_{\mathcal{H}}$ ，由此，MMD 的经验性估计为：

$$\begin{aligned} MMD^2(\mathcal{D}_s, \mathcal{D}_t) &= \frac{1}{m^2} \sum_{i,j=1}^m k(x_i^s, x_j^s) + \frac{1}{n^2} \sum_{i,j=1}^n k(x_i^t, x_j^t) \\ &\quad - \frac{2}{mn} \sum_{i,j=1}^{m,n} k(x_i^s, x_j^t) \end{aligned} \quad \dots (2.20)$$

核函数 k 一般被定义为 U 个基础核函数的凸组合：

$$k(x^s, x^t) = \sum_{u=1}^U \beta_u k_u(x^s, x^t), s.t. \beta_u \geq 0, \sum_{u=1}^U \beta_u = 1 \quad \dots (2.21)$$

本方法选择高斯核函数作为基础核函数。

为了在神经网络中用到 mini-batch 随机梯度下降的方法以及在 CNN 中花费更少的计算时间，Gretton 等 (2012) 提出了一种无偏估计的方法，转化成和的形式，减少到线性的复杂度。给定一个四元组 $z_i = (x_{2i-1}^s, x_{2i}^s, x_{2i-1}^t, x_{2i}^t)$ ，假设 $m = n$ ，MMD 的平方可以写成以下形式：

$$MMD^2(\mathcal{D}_s, \mathcal{D}_t) = \frac{2}{m} \sum_{i=1}^{m/2} h(z_i) \quad \dots (2.22)$$

其中

$$\begin{aligned} h(z_i) = & k(x_{2i-1}^s, x_{2i}^s) + k(x_{2i-1}^t, x_{2i}^t) \\ & - k(x_{2i-1}^s, x_{2i}^t) - k(x_{2i}^s, x_{2i-1}^t) \end{aligned} \quad \dots (2.23)$$

基于域适应的传递式多源迁移框架根据以上理论基础提出。根据上一节中得到的传递式迁移网络 $N(S^i * \dots * S^n)$ ，假设已知前 k 层的特征直接迁移到目标任务 T 上都能取得很好的效果，而 $k+1 \sim L$ 层特征的一般性有明显的下降，迁移性变差，则网络 $N(S^i * \dots * S^n)$ 的前 k 层称为现成层 (off-the-shelf layers)，这些层的特征能够直接用来表示目标数据 T ； $k+1 \sim L$ 层称作适应层，需要通过域适应的方法来提高在目标任务 T 上的特征表达能力。

在一般的域适应方法中 (Long 等, 2015, 2017; Tzeng 等, 2014; Ganin 和 Lempitsky, 2015)，源域数据和目标域数据有着相同的类别，只有目标域数据的标签未知，通过域适应使得二者在某个特征空间的分布更加接近，因此目标域数据的分类结果可以直接用这个空间的特征和源域数据的分类器获得。但在本节的应用与探究过程中，源域任务和目标域任务通常不是针对相同类别标签的分类，因此本节提出了两种框架：一体式域适应迁移和独立式域适应迁移。

一体式域适应迁移框架 (ITL) 将分类和域适应联系起来，如图 2.20(a) 所示，对于需要进行域适应的源数据 S^n 和目标数据 T ，每次随机挑选四个样本 $\{x_{2i-1}^s, x_{2i}^s, x_{2i-1}^t, x_{2i}^t\}$ 作为网络的输入。对于适应层来说，第 l 层的迁移损失函数可以由式(2.22)和(2.23)得到，记为 $mmd_l(x_{2i-1}^s, x_{2i}^s, x_{2i-1}^t, x_{2i}^t)$ ，其中 $l = k+1, \dots, L$ 。目标数据 \mathcal{D}_s 的分类损失函数记为 $\mathcal{L}_C(x_{2i-1}^t, y_{2i-1}^t; x_{2i}^t, y_{2i}^t; \theta_C)$ ，整个网络通过最小化下面的损失函数进行训练

$$\begin{aligned} & \mathcal{L}_C(x_{2i-1}^t, y_{2i-1}^t; x_{2i}^t, y_{2i}^t; \theta_C) \\ & + \lambda \sum_{l=k+1}^L \alpha_l mmd_l(x_{2i-1}^s, x_{2i}^s, x_{2i-1}^t, x_{2i}^t) \end{aligned} \quad \dots (2.24)$$

其中 λ 为权衡迁移损失和分类损失的参数， α_l 表示每个适应层迁移损失函数的权重。在 ITL 框架中，适应层的迁移损失作为分类时的正则项，用一个动态参数 λ 来维持二者之间的平衡，特别是在训练后期 λ 最好乘以 0.1 的衰减率以保证网络在学习的时候主要关注类别的准确性。 α 的值与每层特征的迁移性有关。一般情况下现成层的学习率要小于适应层的学习率，而分类层的学习率要大于适应层的学习率。

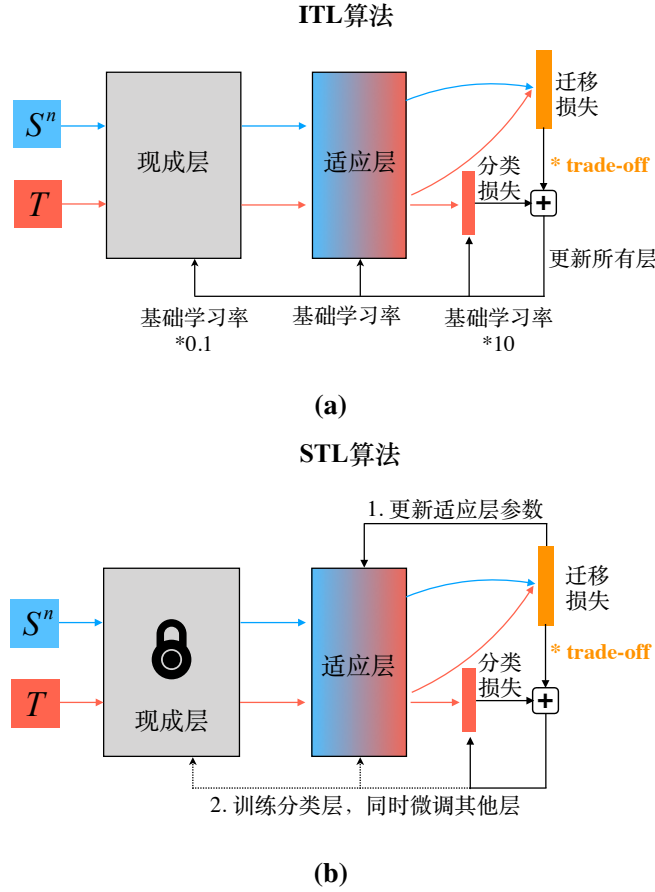


图 2.20 一体式域适应迁移框架 (ITL) 和独立式域适应迁移框架 (STL) 示意图。(a) ITL, (b) STL.

Figure 2.20 The descriptions of integrated transfer learning and separated transfer learning framework. (a) ITL, (b) STL.

同时, 考虑到迁移损失和分类损失在线性结合优化网络的过程中相互制约的问题, 本小节还提出了独立式域适应迁移框架 (STL), 如图 2.20(b) 所示。STL 由两部分构成, 由于现成层具有较强的针对目标域数据的特征提取能力, 首先冻结现成层只借助迁移损失训练适应层, 旨在减小适应层的特征差异, 同时降低了计算成本。然后再将迁移损失和分类损失结合起来训练分类层, 在这个过程中现成层需要采用一个更小的学习率和 λ 。

2.5.2 实验分析

图 2.21 将 MSTAR 和 OpenSARShip 数据在 $H_Net(M)$ 和 $H_Net(S_r * M)$ 第四和第五个卷积层的特征可视化, 可见两种数据在 $H_Net(M)$ 高层的特征分布较为相似, 但在 $H_Net(S_r * M)$ 的第五层特征分布表现出较大的差异。在

$AlexNet_Conv(I * M)$ 和 $AlexNet_Conv(I * S * M)$ 预训练模型中, 前四层特征都表现出对 OpenSARShip 数据集较强的一般性, 但在第五层的时候表现出极强的不适应性。通过可视化 MSTAR 和 OpenSARShip 数据在 $AlexNet_Conv(I * S * M)$ 的第四层和第五层的特征分布状况, 如图 2.21所示, 可以观察到两种数据集在 $AlexNet_Conv(I * S * M)$ 第四层的特征分布差异较小, 在第五层中的特征分布差异较大, 很容易区分。为了使源域和目标域的数据在预训练模型高层的特征分布尽可能相似, 本节实验选取了 $AlexNet_Conv(I * M)$ 、 $AlexNet_Conv(I * S * M)$ 和 $H_Net(S_r * M)$ 三个预训练模型进行验证。

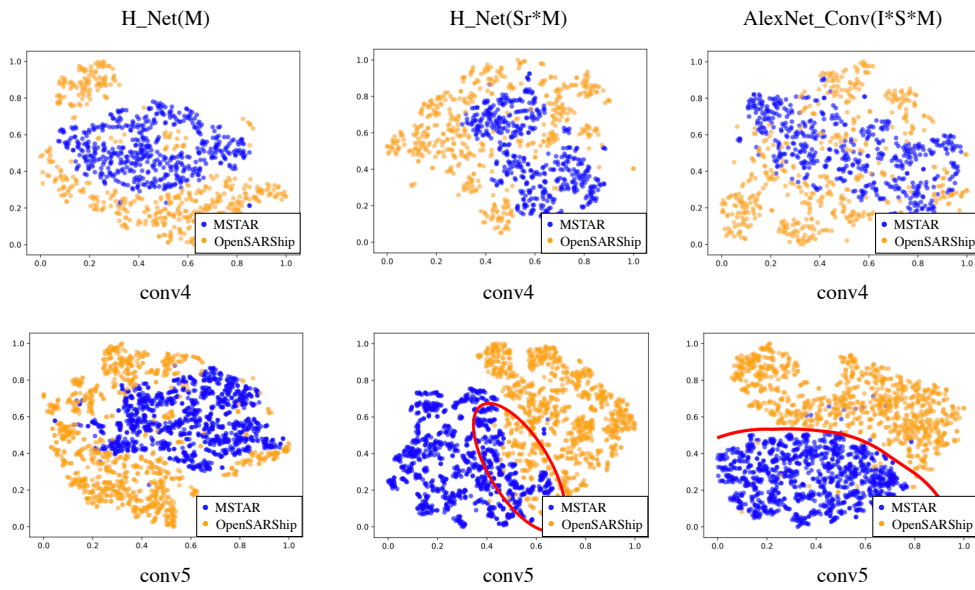


图 2.21 MSTAR 和 OpenSARShip 数据在不同预训练模型的第四层和第五层特征的分布情况

Figure 2.21 The visualization for feature distributions of MSTAR and OpenSARShip data in conv4 and conv5 layers of different pre-trained models.

对于独立式迁移框架 (STL), 根据之前的分析将 $AlexNet_Conv(I * M)$ 和 $AlexNet_Conv(I * S * M)$ 的前四层作为现成层, 第五层作为适应层。在借助域适应更新了适应层之后, 从图 2.22可以观察到 MSTAR 和 OpenSARShip 在第五层的特征分布差异得到了改善。在 STL 的域适应过程中固定了现成层, 使得本身具有良好迁移能力的前四层特征仍然保持其性能。作为一个无监督学习的过程, 意味着可以将无标签的测试集数据加入到训练过程中, 使得适应层对未知数据的鲁棒性更强。接下来结合分类损失和迁移损失用带标签的 OpenSARShip 训练集数据训练分类层, 在这个过程中, 前五个卷积层以 10^{-4} 的学习率仅微调参数,

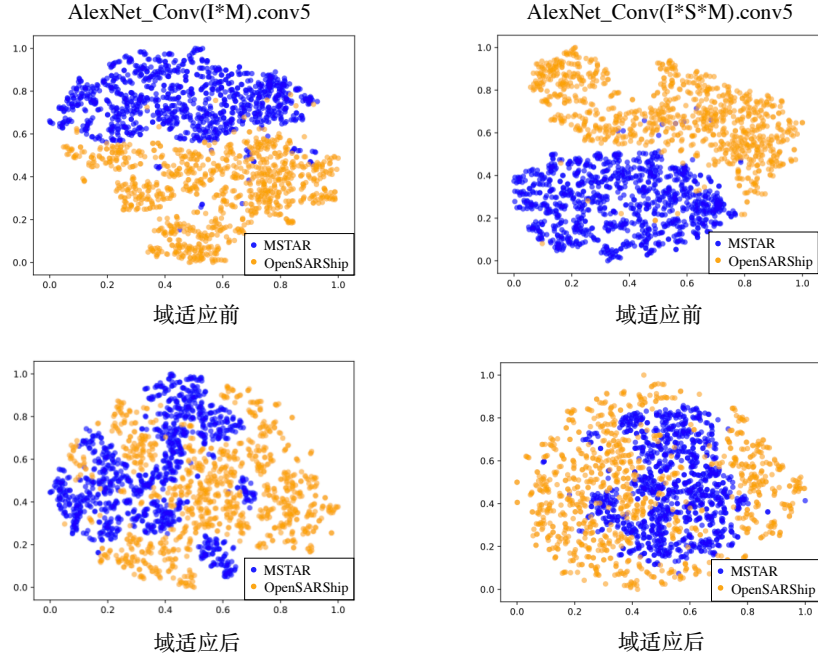


图 2.22 MSTAR 和 OpenSARShip 在 AlexNet_Conv 两种预训练模型第五层的特征域适应前后的可视化效果

Figure 2.22 The visualization for features of MSTAR and OpenSARShip data in the 5th layer of AlexNet_Conv.

以 10^{-2} 的学习率训练分类层。在迁移损失的约束下，网络在拟合分类器的同时也保持了高层特征在两种领域的数据上的较小差异性。在 AlexNet_Conv 情况下 λ 设为 1.5。

表2.12展示了两种迁移框架算法的表现，相比于直接在预训练网络上微调，STL 将 $AlexNet_Conv(I*M)$ 和 $AlexNet_Conv(I*S*M)$ 的效果分别提升了 1.44% 和 1.96%；ITL 算法的表现虽不如 STL，但仍将结果提升了 0.57% 和 0.48%。由于 $AlexNet_Conv(I*M)$ 和 $AlexNet_Conv(I*S*M)$ 两个模型的前四层特征已经具有很强的迁移能力，在改善适应层时保持现成层不变能取得更好的结果，这也表明了分析迁移性对于选择更适合域适应算法的必要性。

对于 $H_Net(S_r * M)$ ，实验表明这里的结果对参数 λ 非常敏感，在将迁移损失和分类损失结合起来的步骤中，需要将 λ 设置为一个更小的值来约束迁移损失在学习过程中的作用。采用域适应方法后的结果如表2.12所示，ITL 和 STL 算法对结果的提升分别为 1.44% 和 0.87%，ITL 算法占优。原因在于 $H_Net(S_r * M)$ 模型第五层特征的迁移性和前四层相比没有极大幅度的下降，前四层特征的微调还存在一定空间，因此在 $H_Net(S_r * M)$ 情况下更有效地选择 ITL 方法进行迁

移。

表 2.12 ITL 和 STL 两种算法相比于微调的性能提升

Table 2.12 The performance of ITL and STL algorithms compared with fine-tuning.

网络	AlexNet_Conv		H_Net
源任务	$I * M$	$I * S * M$	$S_r * M$
微调	89.31%	89.94%	88.44%
ITL	89.88%	90.46%	89.88%
STL	90.75%	91.9%	89.31%

2.6 本章小结

本章以提取 SAR 幅度图像的空间特征为目的，以迁移学习为核心，深入讨论了在 SAR 图像领域中应用卷积神经网络迁移学习的关键性因素，提出适用于 SAR 图像分类的迁移学习算法。以 OpenSARShip 目标识别作为主要研究对象，对基于不同迁移源的预训练网络的各层特征开展了量化的迁移性分析，讨论了数据差异、任务差异和网络差异对预训练网络特征的影响。针对 SAR 目标识别任务，提出基于无标签 SAR 场景图像的迁移学习方法，利用易于获取的大规模 SAR 图像数据作为迁移源解决 SAR 目标样本有限的问题。为了提升基于网络中低层特征的迁移能力，改善高层特征的特殊性，提出基于多种数据源的传递式迁移学习算法和域适应迁移学习方法，在 SAR 图像的相关应用中具有实际应用价值。

本章的主要工作已成文发表：

- HUANG Z, PAN Z, LEI B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data [J]. **Remote Sensing**, 2017, 9(9):907.
- HUANG Z, PAN Z, LEI B. What, where, and how to transfer in SAR target recognition [J], **IEEE Transactions on Geoscience and Remote Sensing**, 2020, 58(4):2324-2336.
- HUANG Z, DUMITRU C O, PAN Z, LEI B, DATCU M. Can a deep network understand the land cover across sensors? [C] **IEEE International Geoscience and Remote**

Sensing Symposium (IGARSS). 2019:9847-9850.

第3章 幅度图像预训练模型

3.1 引言

SAR 图像的土地覆盖和使用分类 (LULC) 是 SAR 图像解译中的一个重要环节, 具有广泛的实际应用价值, 例如农作物分类、植被分类、城市分析、海冰监测等。在基于深度学习的方法中, 由于每个任务的独立性和所使用数据的特殊性, 很多研究针对特定任务和相关的 SAR 图像数据, 设计专门的神经网络来进行学习, 例如Hütt 等 (2016); Zhao 等 (2017b)。在自然图像或光学遥感图像的分类应用中多采取光学预训练网络进行迁移学习, 但上一章的分析表明, 光学预训练模型对 SAR 图像特征的泛化能力随着网络层数的加深急剧恶化, 而依靠大量无标签 SAR 图像进行无监督学习所得的预训练模型对 SAR 图像特征的表示能力较为有限。因此, 本章依托德国宇航中心 (DLR) 提供的 TerraSAR-X 大规模地物分类数据集, 结合上一章内容中所提出的传递式迁移学习方法, 基于残差网络骨架结构训练一个针对 SAR 幅度图像的基准预训练模型。实验表明, 得到的 SAR 图像预训练模型在不同的 SAR 图像分类任务上具有良好的泛化能力, 可以推广应用到其他的基于幅度信息的 SAR 图像应用中。

本章在第二节详细介绍了 TerraSAR-X 大规模 SAR 图像地物标记数据集, 以及存在的三大解译难点, 第三节针对三大难点给出了解决方案和训练方法, 第四节给出了实验结果和分析, 最后一节为本章总结。

3.2 TerraSAR-X 地物分类数据集

本章采用的大规模 SAR 图像地物标记数据集 (Dumitru 等, 2016) 由德国宇航中心 (DLR) 提供, 来自 TerraSAR-X 卫星的聚束模式, 选取极化方式为 HH、分辨率 2.9 米、像素间距 1.25 米的多视地距图像 (Multi-looked Grounded Detected, MGD), 入射角涵盖 20° 到 50° 的范围, 包括升轨和降轨两种轨道模式。数据集中的 SAR 图像切片来自 288 景 TerraSAR-X 图像, 覆盖全球两百多个城市和地区, 以洲为单位的地理分布如图3.1左侧所示。数据集根据景号和地理位置共分成 46 个集合, 每景 SAR 图像按 160×160 像素大小进行无重叠式切片, 每张 SAR 图像切片对应 200×200 平方米的土地, 以保证其中包含至少一种典型的地物类

型。根据每个 SAR 图像切片中主要的地物类别打上对应标签，少部分切片具有两种或三种地物类型标记。数据集的标记共分为三个层级，对应的语义粒度逐渐精细化，例如第一层标记的居所（Settlements）类别对应的第二层标记包括居民区（Residential areas）、商业区（Commercial areas）等，而居民区类别在第三层的标记中又分为高密度居民区（High-density residential areas）、低密度居民区（Low-density residential areas）等。经过数据清洗共统计出七万多张切片，涵盖第一层标记九类、第二层标记 31 类、第三层标记 20 类。由于并非所有数据都有三层的标签，且在第二和第三层的标记中，大部分类别只包含几张或数十张 SAR 切片数据，因此本章只考虑第一层的九类标记，分别是居所 (Settlements)、工业区 (Industrial areas)、公共交通 (Public transportation)、军事设施 (Military facilities)、农田 (Agricultural land)、植被 (Vegetation)、荒地 (Bare ground)、水体 (Water bodies)、未分类 (Unclassified)，其数据量分布如图3.1右图所示。

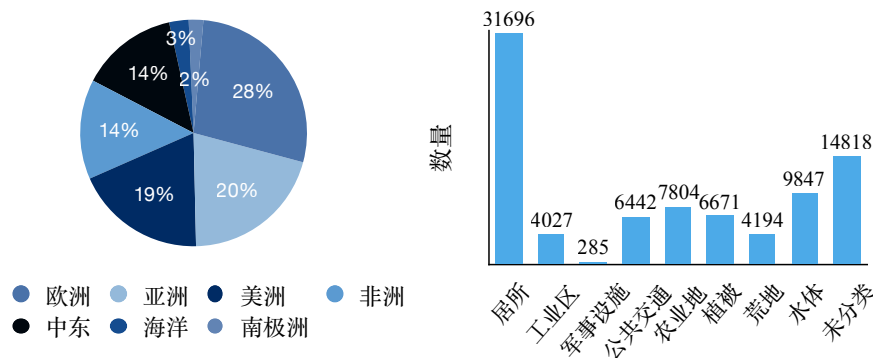


图 3.1 TerraSAR-X 大规模 SAR 地物标记数据集统计信息

Figure 3.1 The statistic information of TerraSAR-X dataset.

对此数据集基于深度学习的自动解译有以下三大方面的难点：

- 数量不均匀分布

虽然整个数据集共包括八万多张 SAR 图像切片，但每类的数据分布非常不均匀。过滤仅具有单个标签的切片进行统计，如图3.1的右图所示，在第一层标注中，居所类别（Settlements）占比最多，数量超过三万张；而军事设施类（Military facilities）只有 285 张切片。考虑到军事设施类别（Military facilities）差距悬殊的少样本量，以及未分类（Unclassified）数据不确定的地物类别，本章只选取了其它七种地物类别进行实验。即便如此，数量最多的居所类仍然是荒地类切片数量的近十倍，这将导致神经网络在学习过程中更偏向于数量较多的类别。

- 丰富的多样性

此数据集覆盖全球多个地区，具有丰富的多样性。例如城市区域的数据，不同国家、地区的城市具有不同的建筑风格和城市规划，由此导致类内差异较大。例如图3.2所示，图中所展示的 SAR 图像切片都属于低密度居民区，但上面的数据来自加拿大的卡尔加里，下面则来自伊拉克的拜亚阿区，二者在视觉上表现出了较大的差异。除此之外，数据集中的切片收集自 2007 年到 2012 年的 TerraSAR-X 卫星数据，在这段时间中，传感器的相关参数（例如辐射校正系数）会发生变化，给不同时期获得的 SAR 图像带来了多样性。

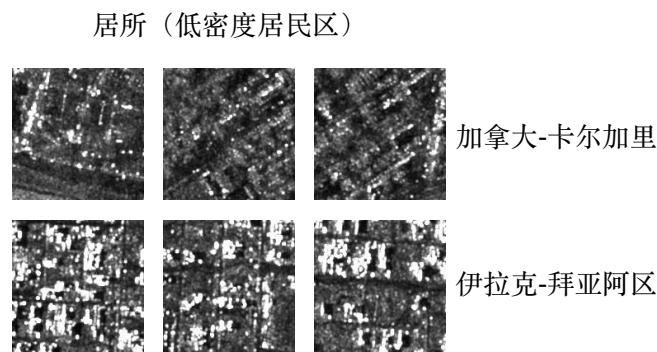


图 3.2 数据集地物多样性示例

Figure 3.2 The diversity example of TerraSAR-X dataset.

- 标签噪声

数据集中的标签噪声主要存在于两个方面：不完全标注或非精确标注。数据集中的每个 SAR 图像切片覆盖的地面范围为 200×200 平方米，根据其中最主要的地物类别进行标记。有的切片会包含多种地物类型，但大多数只赋予了一个标签，此为不完全标注；有的样本标记的类别不是占比最多的地物类型，此为非精确标注。图3.3展示了公共交通和水体类别中的典型样本，以及两个带有噪声标签的样本。噪声标签会导致网络在学习的过程中对非典型样本产生过拟合，从而导致模型的稳健性变差。图3.3就展示了测试过程中因为噪声标签的误导所产生的错误预测结果。

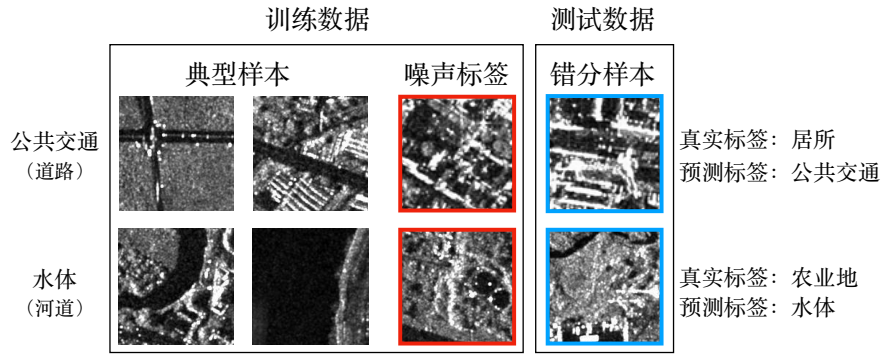


图 3.3 噪声标签示例

Figure 3.3 The noisy labels of TerraSAR-X dataset.

3.3 方法介绍

本章采用数据集第一层标记的七类地物进行分类训练，根据上一节分析的三大难点，本节提出面向 SAR 幅度图像的深度神经网络训练算法。

3.3.1 光学遥感数据集的传递式迁移

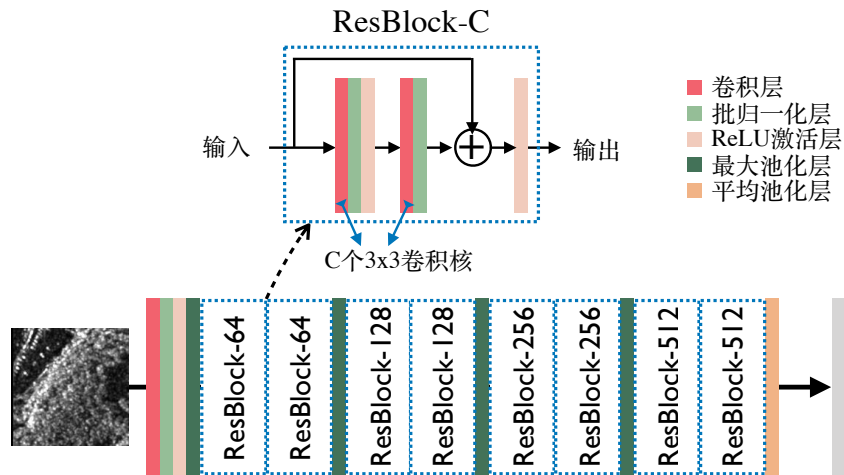


图 3.4 ResNet-18 主干网络示意图

Figure 3.4 The ResNet-18 backbone architecture.

本节选取深度残差网络结构 ResNet-18(He 等, 2016) 作为主干网络, 如图3.4所示, 包含四个重复两次的残差模块 (ResBlock), 每个残差模块由直接映射部分和残差部分组成, 输出由两部分的线性叠加表示。He 等 (2016) 提出的残差结构解决了增加神经网络深度带来的训练退化问题, 是目前最流行的网络设计之一。论文中采用大规模光学自然图像数据集 ImageNet 训练的 ResNet-18 网络具有强

大特征提取能力，根据第二章提出的传递式迁移学习方法，本章提出借助一个和 TerraSAR-X 数据集具有相似标签的光学遥感数据集 NWPU-RESISC45(Cheng 等, 2017) 作为中间任务进行传递式迁移学习，减轻由于自然图像与 SAR 图像之间较大的差异带来的网络迁移性的退化。

NWPU-RESISC45 数据集发布于 2016 年，是一个较新的、规模较大的光学遥感土地覆盖和土地使用数据集，包含 45 类地物类别，共计 31500 张光学遥感图像，每张切片大小为 256×256 像素，收集自谷歌地球遥感数据，分辨率从 0.2 米到 3 米不等。这个数据集也覆盖了全球超过 100 个国家和地区，地物多样性非常丰富。NWPU 数据集和 TerraSAR-X 数据集具有相似的标记类别，二者对比如图 3.5 所示，图中 NWPU 数据集类别后括号中的数字对应 Dumitru 等 (2016) 文中附录表 A.I. 中列出的具体类别。可以看到，本章选取的七类地物在 NWPU 数据集中都有相应的类别。为了进一步缩小高层特征由于数据分布不同而产生的差异，本章根据第二章的研究成果，选取 ResNet-18 网络平均池化之后的特征进行域适应学习。

TerraSAR-X	1-Settlements	2-Industrial areas	3-Military facilities	4-Public transportations	5-Agricultural land	6-Vegetation	7-Bare ground	8-Water bodies	9-Unclassified
数量	31696	4027	285	6442	7804	6671	4194	9847	14818
示例									
NWPU-RESISC45	baseball diamond(123), basketball court(123), church(121), commercial area(122), dense residential area(114), ground track field(123), medium residential(113), mobile home park(11), palace(12), parking lot(131), sparse residential(111), stadium(14), tennis court(123)	industrial areas(2), storage tank(214), thermal power station(21)	None	airplane(421), airport(41), bridge(45), freeway(432), harbor(46), intersection(43), overpass(32), railway station(44), roundabout(434), runway(431), ship(472)	circular farmland(5), rectangular farmland(5), terrace(5)	chaparral, forest(61,62,63,64)	desert(74), mountain(77), meadow(71)	beach(86), lake(82), river(81), sea ice(84)	snow berg, cloud, wetland, golf course, island
示例			None						

图 3.5 TerraSAR-X 数据集与 NWPU-RESISC45 数据集 (Cheng 等, 2017) 各类别标签对比与图像示例

Figure 3.5 The comparison between TerraSAR-X dataset and NWPU-RESISC45 dataset.

3.3.2 重采样和类别加权

考虑到 TerraSAR-X 数据集各类别数量分布不均匀，本章采用了过采样和下采样的策略来保证输入神经网络的每一批数据在每个类别上的均衡性，同时采用 Cost-sensitive 的学习方法 (Kukar 等, 1998)，在样本较少的类别中加大惩罚力度，使网络更多地注意数量较少的样本。对于类别 y 中的样本，计算损失函数时

添加的惩罚因子为

$$w_y = \frac{1}{n-1} \left(1 - \frac{N_y}{\sum_{i=1}^n N_i}\right), \quad \dots (3.1)$$

其中 n 代表类别总数, N_i 表示第 i 类的样本数。

采用不同的评价指标, 例如 Top-1 准确率、Top-2 准确率、查全率、召回率和 F1 分数。F1 分数定义为

$$F_1 = \frac{2}{precision^{-1} + recall^{-1}}, \quad \dots (3.2)$$

其中查全率 $precision = \frac{TP}{TP+FP}$, 召回率 $recall = \frac{TP}{TP+FN}$, TP 、 FP 、 FN 分别表示真正例、假正例和假反例。

3.3.3 联合损失函数

为了减轻标签噪声的影响, 本章采用 Top-k 平滑损失函数 (Berrada 等, 2018) 作为网络的优化目标之一。考虑到本节中只针对第一层级的七种类别, 因此我们主要关注最可能的前两个预测值, 即 Top-2 预测 $y \in \mathbb{R}^2$, 带温度参数的 Top-2 平滑损失函数可以写成如下形式:

$$\begin{aligned} L_{2,\tau}(s, y) = & \tau \log \left[\sum_{y \in \mathcal{Y}^{(2)}} \exp\left(\frac{1}{\tau}(\Delta_k(y, y) + \frac{1}{2} \sum_{j \in y} s_j)\right) \right] \\ & - \tau \log \left[\sum_{y \in \mathcal{Y}_y^{(2)}} \exp\left(\frac{1}{2\tau} \sum_{j \in y} s_j\right) \right]. \end{aligned} \quad \dots (3.3)$$

其中 $s \in \mathbb{R}^n$ 和 y 分别表示在每个类别上的分数向量和真实标签。 τ 表示温度参数, $\Delta_k(y, y) \triangleq \mathbb{1}(y \notin y)$ 表示余量, 具体含义可参考支撑向量机中的余量参数。Top-2 平滑损失函数考虑了网络前两名的预测值, 与只考虑网络排名第一的预测值相比, 一定程度上避免了由于标记不准确导致的网络朝噪声标签上的过拟合。在本方法中, Top-2 平滑损失函数与通用的交叉熵损失函数相结合, 用一个权衡参数 λ 来控制, 再添加模型参数的 L2 正则项, 总的损失函数写为如下形式:

$$\begin{aligned} L_{2,\tau,\mu,\lambda}(s, y) = & (1 - \lambda) \left(- \sum_i s_i \log \frac{\exp(s_i)}{\sum_{j=1}^C \exp(s_j)} \right) \\ & + w_y \lambda L_{2,\tau}(s, y) + \mu \sum_i \|\beta_i\|^2, \end{aligned} \quad \dots (3.4)$$

其中 β_i 表示每层的权重, λ 、 τ 和 μ 分别表示交叉熵损失函数和 Top-2 平滑损失函数之间的权衡参数、温度参数和正则项的权重。

表 3.1 TerraSAR-X 数据集训练集和测试集划分

Table 3.1 The training and testing splitting of TerraSAR-X dataset.

类别	居所	工业区	公共交通	农业地	植被	荒地	水体	总计
训练集	24,930	2,979	4,485	6,029	4,911	2,240	6,826	52,400
测试集一	200	200	200	200	200	200	200	1,400
测试集二	3,155	563	928	925	1,383	1,319	2,133	10,406

3.4 实验结果与分析

首先介绍本实验中所用的训练集和测试集划分。数据集共归为 46 个集合，每个集合的图像具有类似的数据分布。前 39 个集合共包括 53800 张 SAR 图像切片，在每类样本中随机选取 200 个作为测试集一，其余作为训练集；后 7 个集合共包括 10406 张 SAR 图像切片，作为测试集二。测试集一与训练集具有相同的数据分布，但测试集二与训练集和测试集一的分布都不相同，将其作为全新的未知数据。具体的划分情况如表3.1所示。

接下来对 SAR 图像切片进行预处理。数据集中的 SAR 图像切片为多视地距幅度图像，以 16 位整型数据格式存储，绝大多数像素值集中在直方图的最左边，具有均值小方差大的特点。经传统的归一化预处理后直接输入预训练网络，其过大的动态范围会造成训练过程中出现梯度爆炸的现象。因此在预处理过程中加入对数变换降低动态范围，再经过均值方差归一化之后作为网络输入。

几个超参数 λ 、 τ 和 μ 分别设置为 0.2, 1.0 和 0.25，本实验用 Adam 优化策略，在使用 Top-2 平滑损失时初始学习率设置为 10^{-5} ，其他情况下设置为 10^{-4} 。

本节在训练集中每类再划分出 200 例样本作为验证集进行以下几组实验，并记录平均停止训练时间、交叉熵损失函数的变化情况，以及在测试集一上的全局分类准确率，以此来说明本章采用的传递式迁移学习方法的优势：

- 1) RD-TSX: ResNet-18 网络所有层采用随机初始化，在 TerraSAR-X 数据集上从头训练；
- 2) TL-ImageNet2TSX: 采用 ImageNet 预训练的 ResNet-18 模型进行迁移，在 TerraSAR-X 数据集上进行微调；
- 3) TL-ImageNet2RS2TSX: 采用传递式迁移学习方法将 NWPU 数据集作为中间迁移源，再迁移到 TerraSAR-X 数据集上。

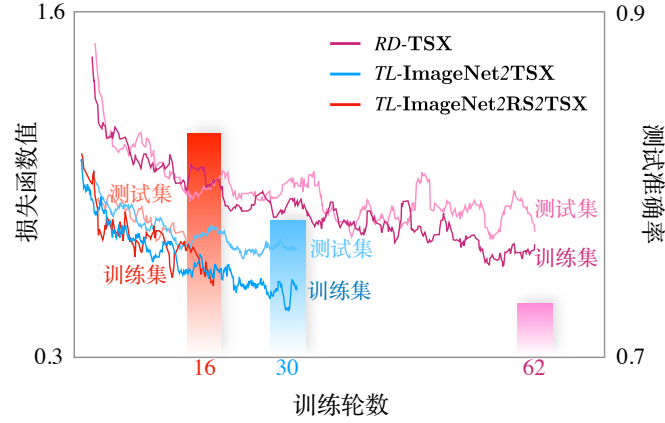


图 3.6 三种训练方式下停止时间、损失函数变化以及测试准确率对比

Figure 3.6 The comparison of stopping time, loss function, and test accuracy among three different training approaches.

如图3.6所示，用 TerraSAR-X 数据集从头训练 ResNet-18 深度卷积残差网络比较困难且耗时长，共经历了 62 轮训练，在验证集上达到 74% 的准确率。采用 ImageNet 的预训练网络进行迁移能有效缩短训练时间（共经历 30 轮训练）并提升网络的分类性能（准确率 75.64%）。但根据上一章的分析，基于 ImageNet 的预训练网络仅在底层的特征对 SAR 图像表现出良好的迁移性，高层的特征泛化性能急剧下降，选取与目标任务相近的中间任务进行传递式迁移学习能提高自然图像预训练网络的中层特征对 SAR 图像的迁移性。在图3.6中可以观察到，采用传递式迁移学习的 TL-ImageNet2RS2TSX 仅用了 16 轮完成训练，并在验证集上的性能超过了 TL-ImageNet2TSX 和 RD-TSX，能达到 78.17%。

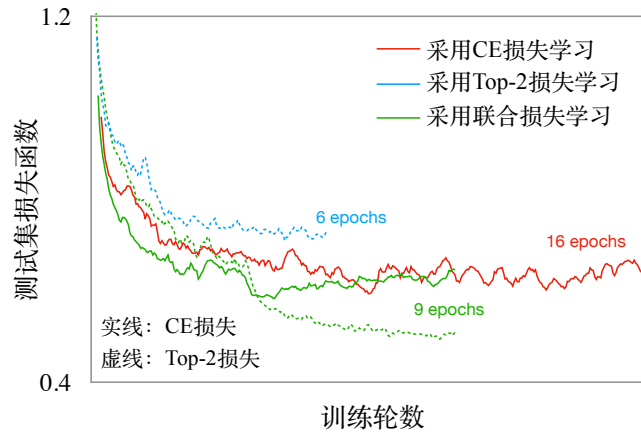


图 3.7 测试集上的各损失函数变化曲线

Figure 3.7 The loss curves on test set.

本章引入 Top-2 平滑损失函数来帮助交叉熵 (CE) 损失在较多噪声标签的情况下更好地优化模型。Top-2 平滑损失考虑了网络预测的前两种地物类型，在当前两个预测标签都不是真值时提供惩罚。图3.7展示了实验中分别采用 CE 损失、Top-2 损失和联合损失作为目标函数进行训练时各损失函数的变化情况，可以观察到 Top-2 平滑损失能加速网络收敛。只采用 Top-2 损失作为优化目标时只通过 6 轮训练网络就已经收敛，而相比之下采用 CE 损失时一共训练了 16 轮。因为 Top-2 平滑损失放宽了对目标函数的约束，使得网络训练速度加快，但这种不严格的约束也会导致网络收敛到一个 Top-2 准确率很高但 Top-1 准确率表现较差的局部最低点，因而只采用 Top-2 损失函数进行优化得到的结果并不好。因此本章选择联合损失函数，将 Top-2 平滑损失作为 CE 损失的约束，这样一来既加快了收敛也提升了网络对于噪声标签的鲁棒性，表3.2的结果表明测试集准确率以及 F1 分数都得到了改善。

表 3.2 在测试集一和测试集二上的测试结果

Table 3.2 The test results of test set 1 and 2.

评价指标 \ 损失函数		CE 损失	Top-2 平滑损失	联合损失
测试集一	F1 分数	0.8318	0.7954	0.8327
	Top-1 准确率	83.14%	79.5%	83.21%
	Top-2 准确率	95.36%	93.29%	95.00%
测试集二	F1 分数	0.709	0.684	0.72
	Top-1 准确率	74.09%	73.24%	76.78%
	Top-2 准确率	91.26%	91.25%	92.62%
训练停止时间		16 轮	6 轮	9 轮

图3.8展示了测试集二中七个集合分别在 CE 损失和联合损失训练下的表现，在大多数新数据上联合损失都具有明显的优势，在平均准确率上整体提升了 2.69%，说明采用联合损失在新的测试集中体现出更强的鲁棒性。图3.9给出了模型在测试集一和二中的混淆矩阵和 Top-2 准确率。

除了设置分布不同的测试集二来验证预训练模型的泛化性能，本节还进行了其他两个实验来进一步说明。图3.10展示了类别激活映射 (Class Activation Map,

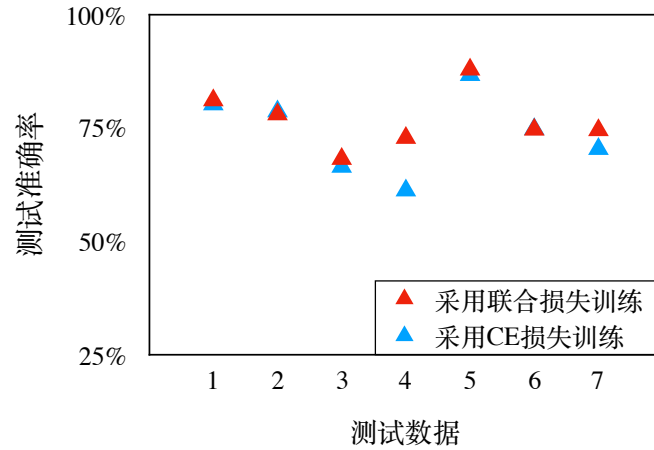


图 3.8 预训练模型在测试集二 7 个集合中分类准确率

Figure 3.8 The test accuracy of the pre-trained model on Test Data 2.

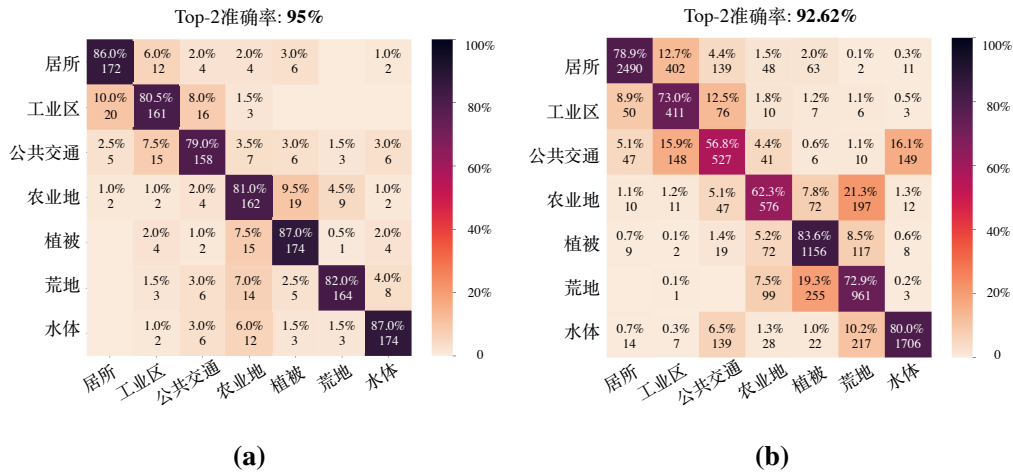


图 3.9 模型在测试集一和测试集二上的混淆矩阵和 Top-2 准确率结果。(a) 测试集一，(b) 测试集二。

Figure 3.9 The test results of Teta Data 1 and Test Data 2. (a) Test Data 1, (b) Test Data 2.

CAM) (Zhou 等, 2016a) 方法对一幅 SAR 场景中的三种类别进行热图分析的结果, 可以观察到在没有额外训练的情况下, 预训练模型所提取的特征在新的 SAR 场景图像中能较好地地区分出居所、工业区及水体等地物。图3.11展示了预训练模型在 10 类 MSTAR 目标识别任务中的迁移结果, 在未经过数据扩充的情况下, 仅通过简单的微调就得到了 99.43% 的准确率, 超过目前的最好结果, 而直接在 ImageNet 预训练的 ResNet-18 模型上进行微调得到的识别率为 98.6%, 相比之下本章提出的模型表现更优异。

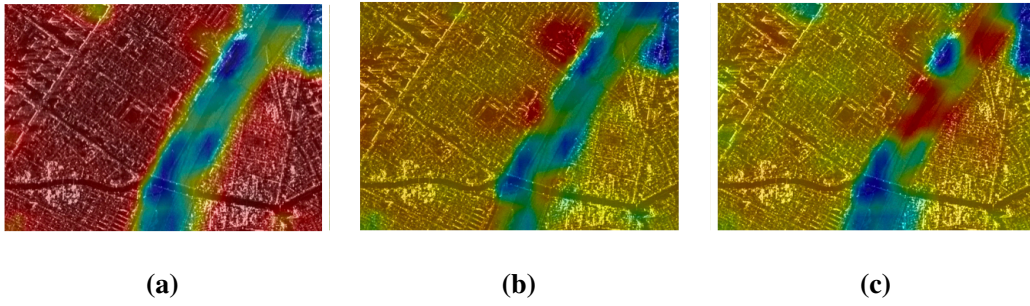


图 3.10 预训练模型利用类别激活映射 (CAM) (Zhou 等, 2016a) 对一景 SAR 图像中几种地物的可视化热度图。(a) 居所, (b) 工业区, (3) 水体。

Figure 3.10 The CAM result with the pre-trained model. (a) Settlements, (b) Industrial areas, (c) Water bodies.

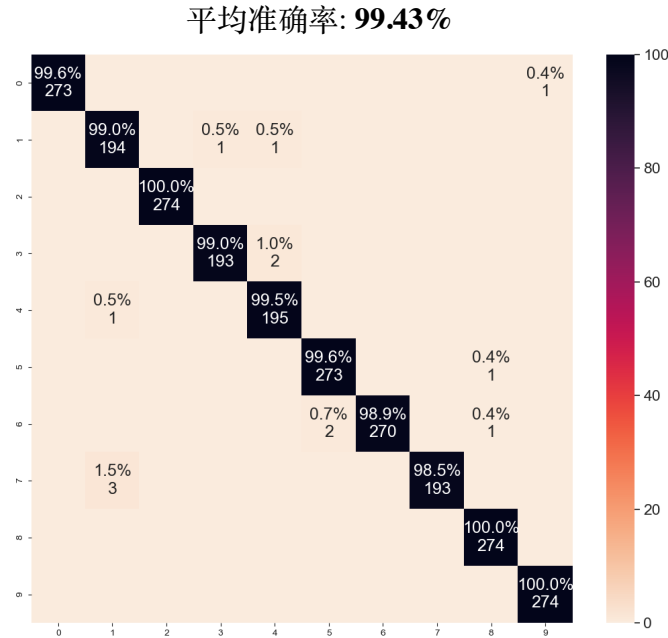


图 3.11 预训练模型迁移至 MSTAR 目标识别任务的测试结果

Figure 3.11 The test result of transferring the pre-trained model to MSTAR target recognition.

3.5 本章小结

本章依托德国宇航中心提供的 TerraSAR-X 大规模地物标记数据集, 以第二章的理论与分析作为研究基础, 基于 ResNet-18 骨架结构和 ImageNet 预训练网络, 以光学遥感数据集作为中间源任务进行传递式迁移学习, 并在最高特征层采用域适应来增强模型在 SAR 图像上的适应能力。提出联合损失函数, 在训练过程中放松对数据标签的约束, 以此克服噪声标签的影响, 提升模型鲁棒性。本章设置多个实验来验证所得深度模型在其他 SAR 图像任务上的泛化能力和特征提

取能力，均取得了很好的效果，实用性较强，未来可作为针对 SAR 幅度图像的预训练网络应用到其他任务中。

本章的主要工作已成文发表：

- **HUANG Z, DUMITRU C O, PAN Z, LEI B, DATCU M.** Classification of large-scale high-resolution SAR images with deep transfer learning [J], **IEEE Geoscience and Remote Sensing Letters**, 2020, Early Access

第4章 面向时频特征的分类方法

4.1 引言

SAR 图像解译的困难很大程度在于主动微波传感器特殊的成像机制导致 SAR 图像表现出的不符合人眼系统的特征。作为主动式侧视雷达系统, SAR 的成像几何属于斜距投影类型, 在原始的 SAR 斜距图像中会出现一些变形, 例如近距压缩、透视收缩、叠掩、阴影等。同时作为一种相干成像传感器, SAR 回波信号具有随机相位, 电磁波的矢量叠加导致 SAR 图像中产生相干斑噪声, 对 SAR 图像的目视判别造成了一定的干扰。SAR 图像中表现出的散射强度和纹理受多种因素的影响: 一方面与 SAR 传感器的参数有关, 例如波长、极化方式、入射角等; 另一方面与目标的物理特性有关, 例如形状、材料、粗糙度等。图4.1(a)展示了 Sentinel-1 卫星在条带模式下对美国休斯敦中心城区的成像, 成像时间均为 2019 年 3 月, 其中左图成像的入射角为 23.34° , 极化方式 VV, 卫星处于降轨, 而中图成像入射角为 32.17° , 极化方式 HH, 卫星处于升轨。可以观察到同一个区域在不同的成像条件下表现出的纹理特性相差较大, 例如红框标记的棒球场和高楼两个人造目标。而不同的目标在 SAR 图像中会呈现相似的纹理, 如图4.1(b)和图4.1(c)所示。

目前大多数针对 SAR 图像分类的深度学习算法将 SAR 数据的幅度, 或经处理后只保留强度信息的产品 (例如 TerraSAR-X 的 MGD 产品、高分三号的二级数据等) 作为输入, 学习图像的空间纹理特征及其与地物语义空间的映射 (Zhao 等, 2017c; Geng 等, 2015, 2017; Chen 等, 2016), 如图4.2中的红框所示。相比之下, 复数 SAR 数据不仅包含地表对雷达波的反射强度, 还包含了相位信息, 通常在极化 SAR 的分析中通过极化散射矩阵反映目标的后向散射特性, 揭示地面目标的物理性质。这些散射特性由复杂的电磁散射物理模型根据雷达工作参数和地物的目标特性来决定, 如图4.2中的蓝框所示。要想实现从已知的 SAR 图像数据中通过理解潜在的物理散射特性推断出地物或目标类别, 需要充分了解这个复杂电磁散射模型的逆过程, 这在实际应用中很难做到。在大量 SAR 图像数据获取的当下, 通过深度神经网络在地物类别与 SAR 图像的后向散射信号以及潜在的散射特性之间建立联系是目前需要解决的问题。

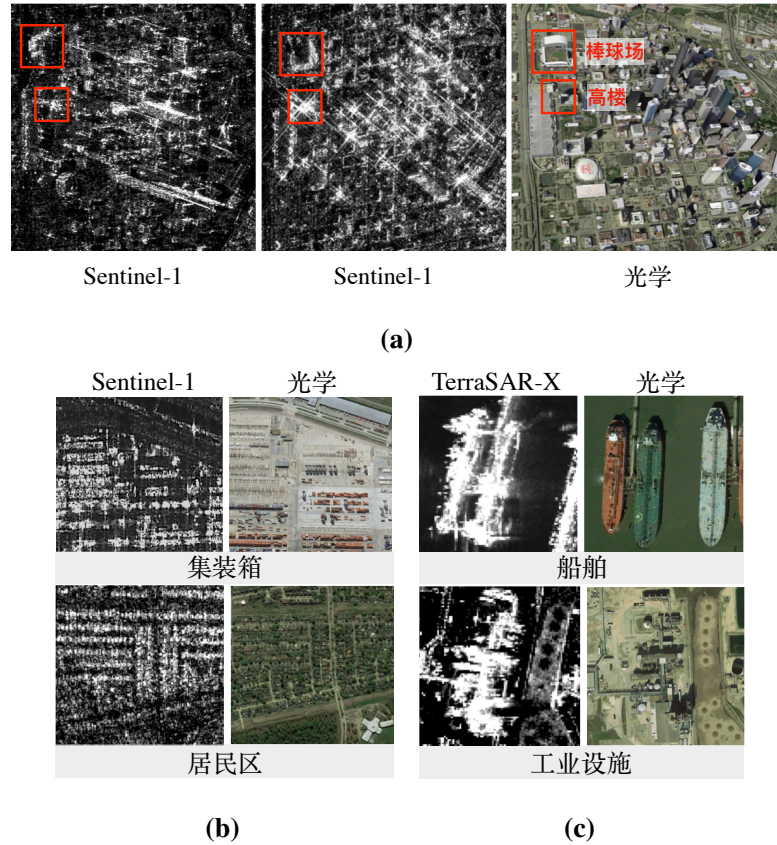


图 4.1 SAR 图像目视解译难点示例。(a) Sentinel-1 卫星对同一个地区的成像对比, (b) 居民区和集装箱的 SAR 图像示例, (c) 船只和工业设施的 SAR 图像示例。

Figure 4.1 The difficulty of visually interpreting SAR images.(a) Two SAR images of the same location by Sentinel-1, (b) Examples of residential area and container, (c) Examples of ship and industry facilities.

理论驱动的物理散射模型和数据驱动的深度学习模型严格意义上属于两个不同的领域,前者基于可靠的物理理论,具有可解释性,但建模较困难,参数难估计;后者善于从大规模数据中自动学习建模,发掘潜在的模式,而在物理意义上的可解释性较弱。Reichstein 等 (2019) 指出,在地球科学中很有必要将物理模型和深度学习模型相结合发挥各自的优势相互补充,这种思路在化学 (Willis 和 von Stosch, 2017) 和大气科学 (Gentine 等, 2018) 等领域已有相关的研究出现。从本章开始的研究内容将跳出传统的基于 SAR 幅度图像的深度学习框架,从 SAR 图像的复数信号出发,利用深度神经网络结合已知的物理模型解决 SAR 图像分类问题。

前面提到仅依靠幅度图像有时难以区分散射较为复杂的场景,如图4.1给出的例子所示。本章利用复数 SAR 图像信号频域带通滤波操作的特殊物理意义,提

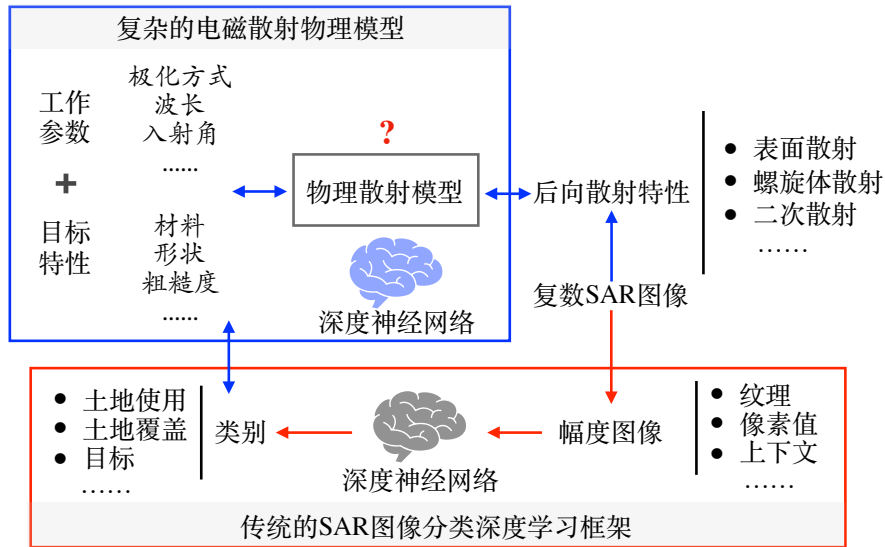


图 4.2 结合散射特性的 SAR 图像分类问题描述

Figure 4.2 Problem description

出利用基于短时傅里叶变换的时频分析方法扩展复数 SAR 图像信号，挖掘 SAR 图像中的非视觉信息用于辅助深度学习模型进行判别。一方面从空间域出发，利用上一章得到的 SAR 图像预训练模型提取初始 SAR 图像空间特征；另一方面从频域出发，设计深度学习自动学习 SAR 图像频域信号的特征，将二者进行融合实现端到端的分类网络搭建。本章依托 Sentinel-1 卫星条带模式下的单视复数 SAR 图像和 TerraSAR-X 卫星聚束模式下的复数 SAR 数据，构建两个 SAR 地物分类复数数据集，再加上 OpenSARShip 船只分类的数据集，验证所提框架在 SAR 图像分类任务中的表现，并设置多个基线模型进行对比。并通过可视化特征分布分析所提方法在 SAR 图像分类中的优势，对具体的样本进行了分析讨论。

本章内容安排如下，第二节介绍复数 SAR 图像二维连续子带分解的过程和物理意义，第三节详细介绍面向复数 SAR 图像时频特征的深度学习分类框架，第四节给出了实验和分析结果，第五节为本章小结。

4.2 SAR 图像的二维连续子带分解

SAR 在成像过程中，传感器沿方位向运动，因此雷达与被成像目标的相对位置在不断地变化。在合成孔径的过程中，雷达将在不同位置接收的目标散射信号进行合成，得到全分辨率的合成孔径图像。在高分辨率 SAR 的情况下，由于雷达波束宽，合成孔径时间长，一些目标在子孔径之间表现出非平稳的特性。早

期的一些研究对复数 SAR 图像在方位向上进行时频分析, 得到不同多普勒频段上的子孔径图像, 每个子孔径图像对应传感器以某个视角注视目标的成像。由于目标的散射特性随方位角变化, 子孔径图像和全孔径图像, 以及不同子孔径图像之间都会有一定的差别。相关研究通过子孔径分解来精确描述目标散射特性随方位视角的变化, 并提出一些应用, 包括利用 SAR 子孔径图像的信息进行动目标检测 (Tupin 和 Tison, 2004)、人造目标或相干目标检测 (Wu 等, 2013; Bovenga 等, 2014), 以及海洋遥感图像分割 (Renga 等, 2019) 等。

另一方面, 高分辨率 SAR 发射带宽很高的线性调频信号以保证距离向的高分辨率, 因此也有一些研究在距离向上进行子带分解, 来获得不同观测频率下的回波散射信号, 用于描述或区分对发射频率敏感的目标 (Ferro-Famil 等, 2005; Bovenga 等, 2014, 2011)。

本小节提出的 SAR 图像二维连续子带分解基于 Singh 和 Datcu (2012) 的研究, 同时在方位向和距离向上进行操作, 并采取经典的短时傅里叶变换时频分析方法。对于一张 SAR 复数图像 $C(x, y)$, 假设二维窗函数 w_B 在距离向和方位向上的带宽分别为 bw_r 和 bw_a , 选取所有可能的中心频率对 (f_r, f_a) 在频域上进行滤波, 得到一系列连续的二维子带图像:

$$S(x, y, f_r, f_a) = \text{FFT}^{-1}[w_B(f_r, f_a) \cdot \text{FFT}(C(x, y))] \quad \dots (4.1)$$

其中 FFT 和 FFT^{-1} 分别表示傅里叶变换和傅里叶逆变换。如此即获得了一个四维的 SAR “超图” (hyper-image) S , x 和 y 两个维度表示空间位置, f_r 和 f_a 两个维度分别表示距离向和方位向的频率坐标。Singh 和 Datcu (2012) 对于四维时频信号 S 提出了六种不同的二维投影来对目标进行目视分析和解译, 本节选取两种极具代表性的投影进行说明。

式4.1计算了给定带宽的所有可能的 SAR 子带图像, 假设固定某个频率对 (f_r, f_a) , 则 $S(x, y, f_r, f_a)$ 表示以 (f_r, f_a) 为中心频率观察到的子带图像, 从视觉上可以认为是原合成孔径图像的低分辨率视图; 若固定某个位置 (x, y) , 则 $S(x, y, f_r, f_a)$ 则表示点 (x, y) 在所有子带图像上的信号, 两个坐标轴分别表示距离向和方位向的频率, 在这里将其称为 SAR 图像在点 (x, y) 的频域信号。图4.3对两个目标切片所生成的四维时频信号 S 的部分子带图像和频域信号进行了幅度可视化。

SAR 图像的二维连续子带分解为 SAR 图像的理解提供了不同的观测角度，下节内容将以此为基础，提出框架的整体流程和具体实现。

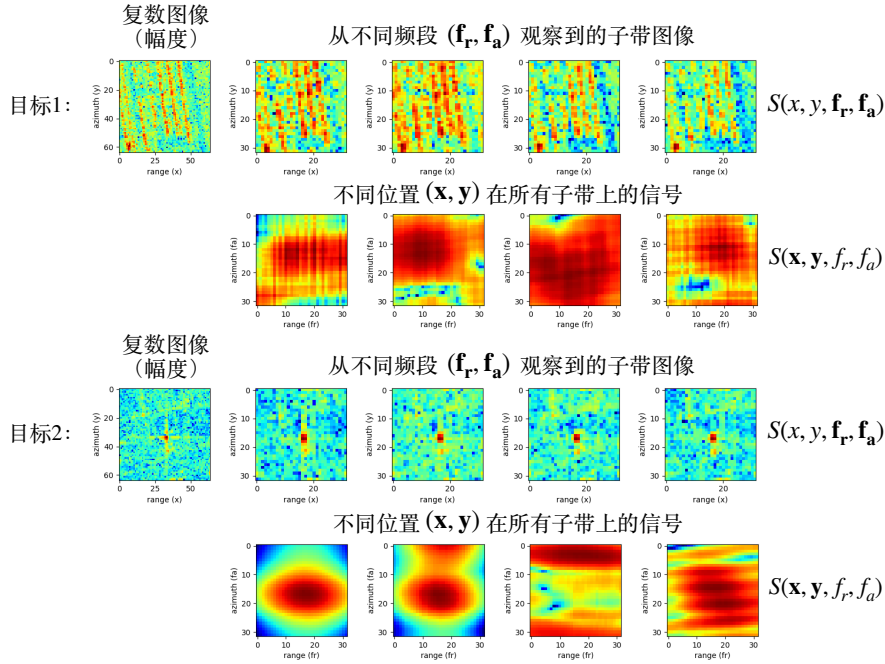


图 4.3 SAR 时频四维信号的理解

Figure 4.3 Different interpretation of 4-D tensor with Time-Frequency Analysis of SAR image.

4.3 面向时频特征的分类方法

本节介绍针对复数 SAR 图像分类的 Deep SAR-Net (DSN) 深度学习框架。

4.3.1 框架总览

本小节首先简单介绍 DSN 框架的主要内容，同时给出必要的符号定义。

将复数 SAR 图像记作 $C(x, y)$ ，其中 x 和 y 分别代表距离向和方位向上的位置。将基于 CNN 的传统 SAR 图像解译算法中采用的地距幅度图像记作 $\mathcal{P}(C(x, y))$ ，其中 \mathcal{P} 表示在单视复数 (SLC) 图像上进行的一系列操作，包括热噪声去除、辐射校正、地理校正、多视等。如图4.2中的红色框内流程所示，传统方法设计一个 CNN 网络用于提取 SAR 图像特征 $\phi(\mathcal{P}(C(x, y)))$ ，再将其映射到地物类别空间 L 中。但过程 \mathcal{P} 只保留了复数图像的幅度信息，同时其中的重采样和滤波等操作虽然使得处理后的 SAR 图像更符合人眼视觉，却丢失了很多信息。基于此，本节提出的 DSN 框架针对复数的 SAR 图像数据，在应用深度学习算法时同时从空间域和频域两个角度去理解信号，使得 SAR 图像的信息能够得

到充分的利用。

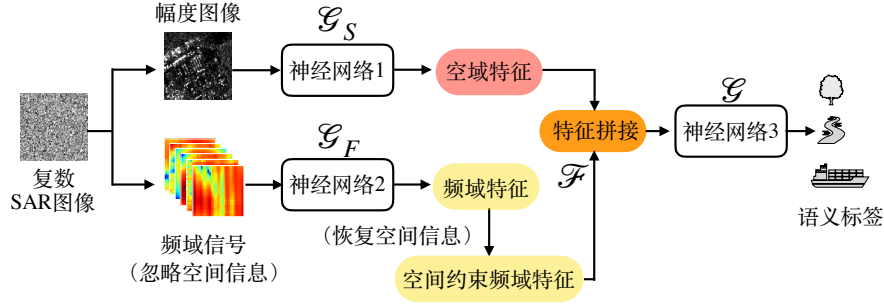


图 4.4 DSN 框架概述

Figure 4.4 DSN workflow

DSN 框架的总体设计如图4.4所示。首先，从全孔径的复数图像 $C(x, y)$ 中得到强度图像 $I(x, y)$ ，用一个传统的 CNN 网络 \mathcal{G}_S 来学习空间域的特征 $\phi_1(x, y)$ ，记作：

$$\phi_1(x, y) = \mathcal{G}_S(I(x, y), \theta_S) \quad \dots (4.2)$$

其中 θ_S 表示 \mathcal{G}_S 中的参数。强度图像 $I(x, y)$ 的边长记为 N_c ，提取的空间域特征 $\phi_1(x, y)$ 维度为 $[C_1, N_1, N_1]$ 。

然后，对于全孔径复数图像 $C(x, y)$ ，按上节内容进行二维连续子带分解得到四维时频信号 S ，并从频域角度观测得到一系列频域信号 $S(\mathbf{x}, \mathbf{y}, f_r, f_a)$ ，忽略其位置信息 (\mathbf{x}, \mathbf{y}) 将其记为 $s(f_r, f_a)_{(\mathbf{x}, \mathbf{y})}$ 。用特征提取器 \mathcal{G}_F 作用在频域信号上，可以得到频域信号的嵌入特征向量 $\phi_2(f_r, f_a) \in \mathbb{R}^n$ ，其长度记为 C_2 ：

$$\phi_2(f_r, f_a) = \mathcal{G}_F(s(f_r, f_a), \theta_F) \quad \dots (4.3)$$

其中 θ_F 是 \mathcal{G}_F 中的参数。在这之后试图恢复 ϕ_1 和 ϕ_2 两种特征在空间上的对应关系，因此将 $\phi_2(f_r, f_a)$ 按照 $s(f_r, f_a)_{(\mathbf{x}, \mathbf{y})}$ 在 $S(\mathbf{x}, \mathbf{y}, f_r, f_a)$ 中的排列堆叠起来得到空域对齐的频域特征 $\Phi(x, y, \phi_2)$ ：

$$\Phi(x, y, \phi_2) = \{\phi_2(f_r, f_a)_{(\mathbf{x}, \mathbf{y})}\}_{\mathbf{x}, \mathbf{y}=1}^{N_c} \quad \dots (4.4)$$

其中 $\Phi(x, y, \phi_2)$ 的维度为 $[N_c, N_c, C_2]$ ，将其进行维度变换后转换成维度为 $[C_2, N_c, N_c]$ 的张量 $\Phi(\phi_2, x, y)$ 。

接下来，将得到的空间域特征 $\phi_1(x, y)$ 和空域对齐的频域特征 $\Phi(x, y, \phi_2)$ 进行融合，得到 $\mathcal{F}(\phi_1, \Phi)$ ，再用函数 \mathcal{G} 将其映射到标签空间 L 中：

$$L = \mathcal{G}(\mathcal{F}(\phi_1, \Phi), \theta) \quad \dots (4.5)$$

其中 θ 是 \mathcal{G} 中的参数。这样便得到了一个将复数 SAR 图像的全孔径幅度信息和子带频域信息映射到标签域上的系统。

上述的几个模块 \mathcal{G}_S 、 \mathcal{G}_F 、 \mathcal{F} 和 \mathcal{G} 组成了本文提出的 DSN 框架，如图4.4所示。下一小节将详细介绍如何借助深度卷积神经网络实现 DSN。

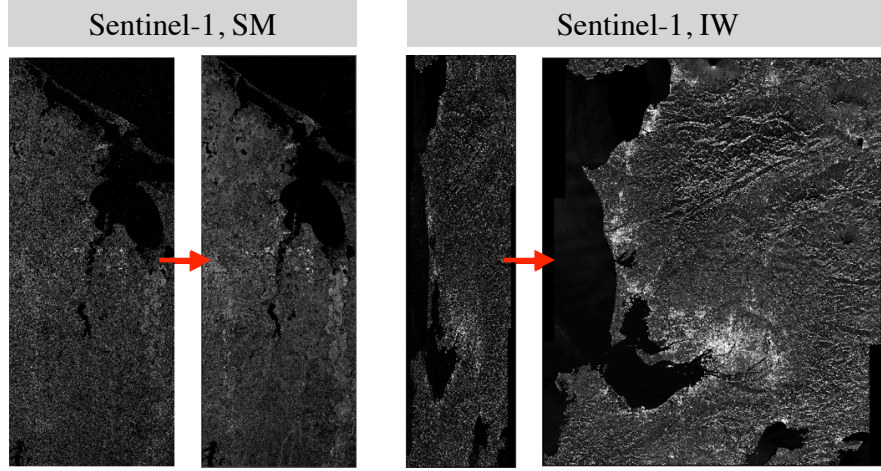


图 4.5 单视复数图像数据的强度图像和多视地距图像对比示意图

Figure 4.5 The comparison between intensity image of SLC product and the MGD image.

4.3.2 全孔径幅度图像的迁移学习

对一个复数的 SAR 图像切片 $C(x, y) = A(x, y) + jB(x, y)$ ，其强度图像由 $I(x, y) = A(x, y)^2 + B(x, y)^2$ 给出。当距离向和方位向上的分辨率和像素间隔相差较小时，多视和斜地转换等操作对图像形变的影响较小， $I(x, y)$ 与地距幅度图像 $\mathcal{P}(C(x, y))$ 在视觉上较为相似。图4.5展示了 Sentinel-1 卫星在不同的成像模式下单视复数 SAR 的强度图像和多视地距图像的对比。左图为条带模式，距离向和方位向分辨率分别为 2.5 米和 3.6 米，像素间隔分别为 2.2 米和 3.5 米；右图为干涉宽幅模式，距离向和方位向分辨率分别为 2.7 米和 22.5 米，像素间隔分别为 2.3 米和 14.1 米。

因此本小节提出基于 TerraSAR-X 多视地距 SAR 图像的迁移学习，将第3章中介绍的用 TerraSAR-X 大规模地物分类数据集训练的深度残差卷积网络作为迁移源，用于本节全孔径 SAR 幅度图像的空间特征提取。基于第3.4小节的探索，此预训练模型的底层和中层特征对于 $I(x, y)$ 有较好的特征表示，因此这里将前四个残差模块作为 \mathcal{G}_S 迁移到 $I(x, y)$ 上，如图4.6的模块一所示，提取出中层特征 $\phi_1(x, y)$ ，最终的特征图共有 128 个通道。一个残差模块的组成如图4.7左侧所示。

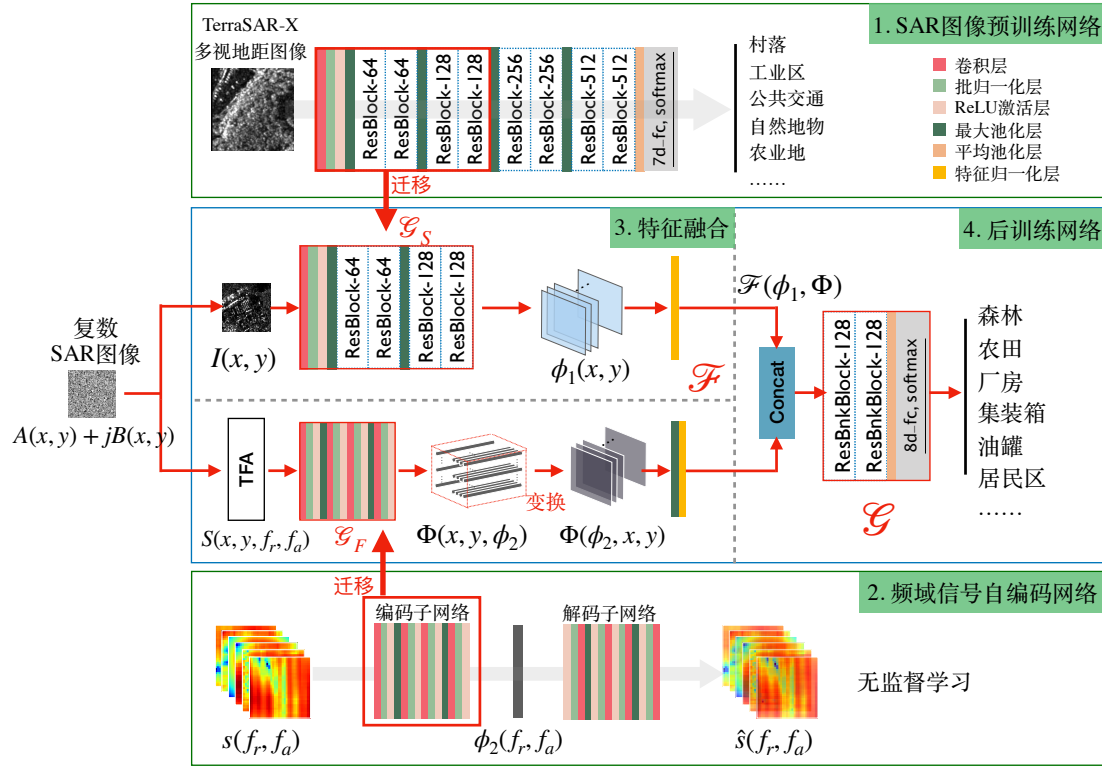


Figure 4.6 The detailed implementations of DSN framework.

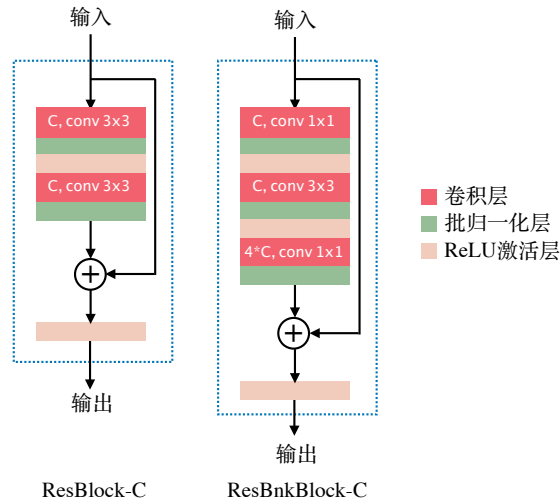


Figure 4.7 The architecture of ResBlock and ResBnkBlock.

4.3.3 频域信号的自编码网络

对一个大小为 $N_c \times N_c$ 的复数 SAR 图像切片 $C(x, y)$ ，其四维时频信号 S 可由式4.1得到，维度为 $[N_x, N_y, N_{f_r}, N_{f_a}]$ 。在这里设置滤波器函数 w 的带宽为合成孔径图像带宽的一半，即 $bw_r = \frac{1}{2}BW_r$ ， $bw_a = \frac{1}{2}BW_a$ ，其中 BW_r 和 BW_a 分别代表合成孔径图像距离向带宽和多普勒带宽。为了使问题进一步简化，在这里设置参数使得 $N_x = N_y = N_{f_r} = N_{f_a} = \frac{1}{2}N_c$ 。

根据上节的分析，对于特定的频率对 (f_r, f_a) ， $S(x, y, f_r, f_a)$ 表示一个子带图像，反映了图像在子视状态 (f_r, f_a) 下的散射表现。对于特定的空间位置 (x, y) ， $S(x, y, f_r, f_a)$ 表示 SAR 图像 $C(x, y)$ 从所有子视状态进行观察时点 (x, y) 对应的频域信号，涵盖了所有频段的信息。DSN 框架对四维时频信号 S 采取先频域再时域的理解方式，忽略 $S(x, y, f_r, f_a)$ 的空间信息，设计一个堆叠式卷积自编码器来学习频域信号 $s(f_r, f_a)$ 潜在的幅度特征 $\phi_2(f_r, f_a) \in \mathbb{R}^n$ ，如图4.6的模块二所示， \mathcal{G}_F 代表卷积自编码器的编码部分， $\tilde{\mathcal{G}}_F$ 代表解码部分。其中编码器的结构如表4.1所示，优化过程通过最小化输入频谱 $s(f_r, f_a)$ 和解码之后的频谱 $\hat{s}(f_r, f_a)$ 的均方根误差来实现：

$$Loss_{mse} = \sum \|\tilde{\mathcal{G}}_F(\mathcal{G}_F(s(f_r, f_a))) - s(f_r, f_a)\|_2 \quad \dots (4.6)$$

有关堆叠式卷积自编码器的具体设计和细节可参照第2.3.1.1小节的内容。

4.3.4 时频特征融合和后续子网络

以 $N_c = 64$ 为例，得到的四维时频信号 $S(x, y, f_r, f_a)$ 维度为 $[32, 32, 32, 32]$ 。强度图像 $I(x, y)$ 通过 TerraSAR-X 的预训练层 \mathcal{G}_S 提取的空间域特征 $\phi_1(x, y)$ 维度为 $[128, 16, 16]$ 。对每个丢弃了空间信息的频域信号 $s(f_r, f_a)$ ，由频域信号自编码器 \mathcal{G}_F 提取的特征 $\phi_2(f_r, f_a)$ 的大小为 $[128,]$ 。当一系列的频域特征 $\phi_2(f_r, f_a)$ 进行空域对齐之后，形成维度为 $[32, 32, 128]$ 的 $\Phi(x, y, \phi_2)$ ，通过维度变换形成 $[128, 32, 32]$ 的张量。接下来通过一个最大池化层对 Φ 进行下采样，归一化之后与 ϕ_1 在空间维度上进行拼接，得到维度为 $[256, 16, 16]$ 的融合特征 $\mathcal{F}(\phi_1, \Phi)$ ，如图4.6的模块三所示。

对融合特征进行后续学习的子网络由两个残差瓶颈模块 (ResBnkBlock) 组成，记作 \mathcal{G} ，将融合的特征 $\mathcal{F}(\phi_1, \Phi)$ 通过 \mathcal{G} 映射到标签域中，如图4.6的模块四所示。这里采用的残差瓶颈模块以 4 为倍数减少了每层特征图的通道数，节省

了大量的计算消耗，模块构成如图4.7所示，具体细节可参考文献 (He 等, 2016)。表4.1列出了几个子网络的卷积核参数设计。总的算法流程如算法1所示。

表 4.1 DSN 框架三个子网络的卷积核设计

Table 4.1 The convolutional kernel design of DSN framework.

$\mathcal{G}_S(\theta_S)$	$\mathcal{G}_F(\theta_F)$	$\mathcal{G}(\theta)$
conv $64 \times 7 \times 7$	conv $32 \times 5 \times 5$	ResBtnBlk-128 ($\times 2$)
ResBlk-64 ($\times 2$)	conv $64 \times 5 \times 5$	conv $128 \times 1 \times 1$
conv $64 \times 3 \times 3$	conv $64 \times 3 \times 3$	conv $128 \times 3 \times 3$
conv $64 \times 3 \times 3$	conv $128 \times 4 \times 4$	conv $512 \times 1 \times 1$
ResBlk-128 ($\times 2$)		
conv $128 \times 3 \times 3$		
conv $128 \times 3 \times 3$		

算法 1 DSN 训练流程

输入: 训练集 $D = \{C_k(x, y), l_k\}_{k=1}^m$; 预训练卷积网络 \mathcal{G}_S ; 窗函数 $w(f_r, f_a)$

输出: 频域信号卷积自编码器 \mathcal{G}_F ; 后续子网络 \mathcal{G}

```

1: function TRAIN_ $\mathcal{G}_F(D, w(f_r, f_a))$ 
2:   训练频域信号自编码网络
3:   for all  $C_k(x, y) \in D$  do
4:     计算当前雷达四维时频图  $S_k(x, y, f_r, f_a) = \text{FFT}^{-1}[w(f_r, f_a) \times \text{FFT}(C_k)](x, y)$ 
5:     丢弃空间信息，得到频域信号集合  $D_k^F = \{s_k(f_r, f_a)_{(x,y)}\}_{x,y=1}^{N_c/2}$ 
6:     for all  $s_k^i(f_r, f_a) \in D_k^F$  do
7:       计算当前频域信号的自编码网络输出  $\hat{s}_k^i(f_r, f_a) = \tilde{\mathcal{G}}_F(\mathcal{G}_F(s_k^i(f_r, f_a)))$ 
8:       计算目标函数  $Loss_{mse} = \sum \|\hat{s}_k^i(f_r, f_a) - s_k^i(f_r, f_a)\|_2$ 
9:       反向传播，更新网络参数  $\theta_F$ 
10:    end for
11:  end for
12: end function
13: function TRAIN_ $\mathcal{G}(D, D^F)$ 
14:   训练后续子网络
15:   for all  $(C_k(x, y), l_k) \in D$  do
16:     计算当前强度图像  $I_k(x, y) = \text{Real}(C_k)^2 + \text{Img}(C_k)^2$ 
17:     计算空间特征  $\phi_1(x, y)$ 
18:     for all  $s_k^i(f_r, f_a) \in D_k^F$  do
19:       计算频域特征  $\phi_2^i(f_r, f_a)$ 

```

```

20:     end for
21:     计算空域对齐的特征  $\Phi(x, y, \phi_2)$ 
22:     下采样  $\Phi$ 
23:     特征归一化  $\phi_1, \Phi$ 
24:     特征融合  $\mathcal{F}(\phi_1, \Phi)$ 
25:     计算当前后续子网络输出  $\hat{l}_k = \mathcal{G}(\mathcal{F}(\phi_1, \Phi))$ 
26:     计算目标函数  $Loss_{cls} = CrossEntropyLoss(l_k, \hat{l}_k)$ 
27:     反向传播, 更新网络参数  $\theta$ 
28: end for
29: end function

```

4.4 实验与分析

4.4.1 数据集介绍

本节共选取三个数据集进行算法验证, 其中两个为自行标注的数据集 S1 和 TSX, 一个为公开数据集 OPS, 下面将依次进行介绍。

S1 数据集来自 Sentinel-1 卫星提供的 SAR 图像。Sentinel-1 卫星由欧洲空间局 (ESA) 主持发射, 工作在双极化 C 波段, 共提供四个成像模式的数据。根据上一小节中的分析, 为了使选取的单视复数 SAR 数据与多视地距幅度图像共享相似的纹理特征, 本节选取了条带模式下由波束三和波束四生成的三张来自三个不同城市的单视复数 SAR 图像, 具有相近的距离向和方位向分辨率及像素间隔。表4.2给出了本次实验数据的一些重要参数。

S1 数据集包括八种类别, 分别是五种以人造目标为主的土地使用类别: 厂房、油罐、集装箱、居民区以及高楼, 和三种自然土地覆盖类别: 森林、农田和水体。对照光学谷歌地球数据, 以 64×64 像素大小为标准在 SAR 图像上进行切片, 根据像素间隔计算 64×64 像素大小的切片在地面上对应 200×200 平方米左右的土地, 基本符合Dumitru 等 (2016) 提出的 TerraSAR-X 数据集的土地覆盖范围。每类地物的数量和幅度图像以及对应光学图像参考图的示意如图4.8所示, 为了实验需要, S1 数据集一共标记了 2550 张切片, 并且具有较均衡的类别分布, 每类大约有 300 张左右的切片, 所有的 SAR 复数切片数据存储为 16 比特的 I/Q 双通道。

TSX 数据集来自 TerraSAR-X 卫星提供的三景 SAR 图像, 采用聚束模式 (SL)

表 4.2 Sentinel-1 条带模式下的 SLC 图像参数

Table 4.2 The parameters of SLC images in Sentinel-1 Stripmap mode.

图像名	BBE7	4312	EBAD
地理位置	休斯敦	芝加哥	圣保罗
极化方式	HH	HH	HH
波束编号	S3	S4	S4
斜距分辨率 (m)	2.5	3	3
方位分辨率 (m)	3.6	4.8	4.8
斜距像素间隔 (m)	2.2	2.6	2.6
方位像素间隔 (m)	3.5	4.1	4.1
入射角 (°)	31.2	36.4	36.4


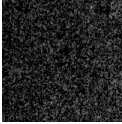
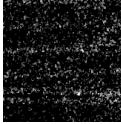
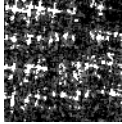
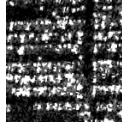
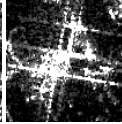
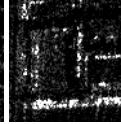
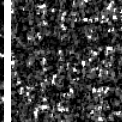
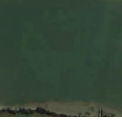



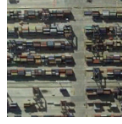

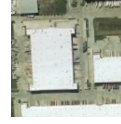

类别	自然地物类				人造目标类			
	水体	森林	农田	油罐	集装箱	高楼	厂房	居民区
SAR 图像 示例								
光学 图像 示例								
数量	350	325	370	342	216	372	223	352

图 4.8 S1 数据集

Figure 4.8 S1 dataset description.

下的单视斜距复数格式 (SSC)，分别位于休斯敦和圣安东尼奥地区，具体参数如表4.3所示。在进行标注之前，SSC 复数图像分别在 I、Q 通道上进行了多视处理，在保留复数数据的基础上将距离向和方位向上的像素间隔统一至 1m。与 Sentinel-1 卫星所提供的 SAR 图像相比，TerraSAR-X 卫星在聚束模式下的成像具有更高的分辨率，为地面目标提供了更清晰的细节。因此，TSX 数据集包含了精细标注的五类人造目标，分别为 352 张油罐切片、341 张铁路切片、225 张船只切片、305 张工业设施切片、以及 390 张居民区切片，切片大小为 64×64 像素，并保证每个切片中含有一个目标主体，如图4.9所示。这个 TSX 精确人造目标数据集将有利于评估所提 DSN 算法在目视解译困难的条件下对于散射复杂的人造目标是否能较好地提取潜在特征并进行区分。

表 4.3 TerraSAR-X 卫星聚束模式下的 SSC 图像参数

Table 4.3 The parameters of SSC images in TerraSAR-X Spotlight mode.

图像名	1550	3142	3527
地理位置	休斯敦	圣安东尼奥	圣安东尼奥
极化方式	HH	VV	VV
斜距分辨率 (m)	0.58	0.58	0.58
方位分辨率 (m)	0.23	0.23	0.23
像素间隔 (m)	1	1	1
入射角 (°)	30	40	26.5

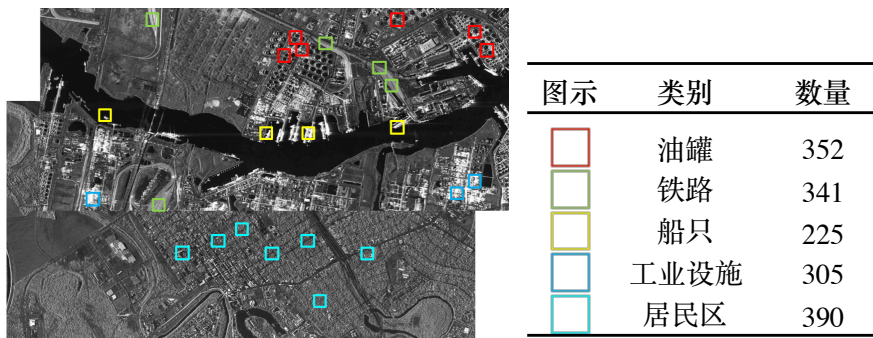


图 4.9 TSX 数据集

Figure 4.9 TSX dataset description.

OPS 数据集来自Huang 等 (2018) 提出的 OpenSARShip 船舶公开数据集, 采用 Sentinel-1 卫星的干涉宽幅模式 (IW) 数据, 包含单视复数图像 (SLC) 和地距 (GRD) 两种格式。本节选取了 VV 极化的 SLC 格式的复数数据来进行实验。原始的 OpenSARShip 数据集涵盖了 17 种类型共计上万的船只, 但每类的数据分布相差较大, 例如货船类 (Cargo) 有 8470 个样本, 而拖船类 (Towing) 只有 4 个样本。参考Huang 等 (2018) 论文中采用的验证数据集, 本节提出的 OPS 数据集由三类主要的船舶类别组成, 分别为散货船 (Bulk Carrier), 集装箱船 (Container Ship) 以及油船 (Tanker), 只选取了大小适中的 SAR 图像切片, 忽略过大或过小的数据, 最终挑选了共 418 张图像, 将切片以目标为中心裁剪为 64×64 的大小。OPS 的每类示例图和数量见图4.10所示。

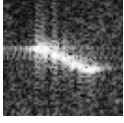

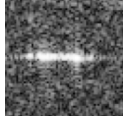



类别	散货船	集装箱船	油船
Sentinel-1 SAR图像			
实际图像			
数量	120	182	116

图 4.10 OPS 数据集

Figure 4.10 OPS dataset description.

4.4.2 实验设置

从 TerraSAR-X 多视地距幅度图像的预训练网络中迁移卷积层时, $\mathcal{G}_S(\theta_S)$ 可作为现成的特征提取器。对于大小为 64×64 的复数 SAR 图像, 频域窗函数采用大小为 32×32 的汉明窗。用频域信号训练堆栈自编码器时, 用于训练的 $s(f_r, f_a)$ 的数量是数据集中复数图像切片 $C(x, y)$ 的 $N_x \times N_y = 1024$ 倍, 在小样本的条件下也能提供大量的训练数据, 因此在训练过程中直接对深层的堆叠结构采用一次训练模式, 区别于 2.3.1.1 小节采用的逐层训练模式。网络训练使用 SGD 优化器, 批大小设置为 200, 初始学习率设置为 0.1, 权重衰减参数 0.0005。在将 \mathcal{G}_S 和 \mathcal{G}_F 进行迁移并训练 \mathcal{G} 的时候, 参数 θ_S 和 θ_F 被固定, 只需对参数 θ 采用 He 式初始化并利用梯度下降更新, 这里采用 Adam 优化器, 初始学习率设置为 0.01。

在验证 DSN 算法的实验中, 每个数据集分别随机挑选 90%、70%、50% 和 30% 的数据进行训练, 剩下的用来测试, 以验证 DSN 算法在少样本条件下的效果。为了增加实验结果的可信度, 每个比例的训练数据实验都进行了五次随机分配训练测试集, 最后对五次实验的结果进行统计, 计算均值和标准差进行比较。算法验证采用多种评价参数, 包括总体准确率 (OA)、F1 分数以及 ROC 曲线等。

本节实验还设置了多个基线模型用于对比 DSN 算法, 具体如下:

- **CNN:** 只采用 SAR 图像幅度信息空间特征的 CNN 网络

采用 DSN 中的 \mathcal{G}_S 和 \mathcal{G} 部分, 将所有的网络参数 θ_S 和 θ 都进行随机初始化, 用强度图像 $I(x, y)$ 从头训练;

- **TL-CNN:** 采用迁移学习的 CNN 网络

采用 DSN 中的 \mathcal{G}_S 和 \mathcal{G} 部分, 网络参数 θ_S 由 TerraSAR-X 预训练网络迁移得到, θ 随机初始化;

- **F-CNN: 只采用频域信号特征的 CNN 网络**

采用 DSN 中的 \mathcal{G}_F 和 \mathcal{G} 部分, 网络参数 θ_F 由预训练自编码网络迁移得到, θ 随机初始化;

- **CV-CNN(Zhang 等, 2017): 复数卷积神经网络**

采用 DSN 中 \mathcal{G}_S 和 \mathcal{G} 的网络结构, 将卷积运算等操作改为在复数域上进行, 直接将复数 SAR 图像作为输入。

需要注意, 在 CV-CNN 基线模型中, 由于复数卷积等操作在实际运算中将实部与虚部分开进行, 因此 CV-CNN 的参数是具有相同结构的实数卷积网络参数的两倍。因此, 为了保证对照实验组有相同的参数量, 本节将 CV-CNN 中每层的特征图数量减为 \mathcal{G}_S 和 \mathcal{G} 的一半。

4.4.3 评价指标分析

对 S1、TSX 和 OPS 三个数据集, 本节对不同比例训练集情况下传统的 CNN 和本章所提 DSN 算法进行了实验, 在测试集上的各评价指标, 如总体平均准确率、F1 分数的均值和误差棒分别如图4.11、图4.12和图4.13所示。总体来看, DSN 与只面向 SAR 幅度图像的 CNN 方法相比展现出明显的优势, 在三个数据集上均优于 CNN 模型。对比图4.11中 CNN 和 TL-CNN 的表现可以看到, 尽管来自不同的传感器、成像模式、产品级别, 但 TerraSAR-X 预训练网络的中低层特征仍然可以较好地迁移到 Sentinel-1 数据上, 并取得良好的效果, 这也佐证了第3章的研究结果。

图4.14给出 S1 数据集在采用 90% 的数据进行训练时, DSN 和 CNN 模型对每个单独类别进行预测的 F1 分数和准确率。可以看到在一些类别上 DSN 相比于 CNN 的优势更加明显, 例如厂房和集装箱; 而在某些类别上 DSN 和 CNN 的表现不相上下, 例如水体、森林和农田, 这些类别都属于自然土地覆盖类别。在这里做出合理推测, 即 DSN 在人造目标类别的分类上相比于只基于空间纹理信息的普通卷积神经网络具有更大的优势, 而在自然类别的分类上优势不明显。为了验证这一推论, 本节将三类自然类别和五类人造目标类别在两种模型上的分类结果通过 ROC 曲线进行可视化, 如图4.15(a) 和4.15b(b) 所示。可以观察到 DSN 在三类自然类别的识别中与 CNN 相比并不占优势, 在人造目标类别中却表

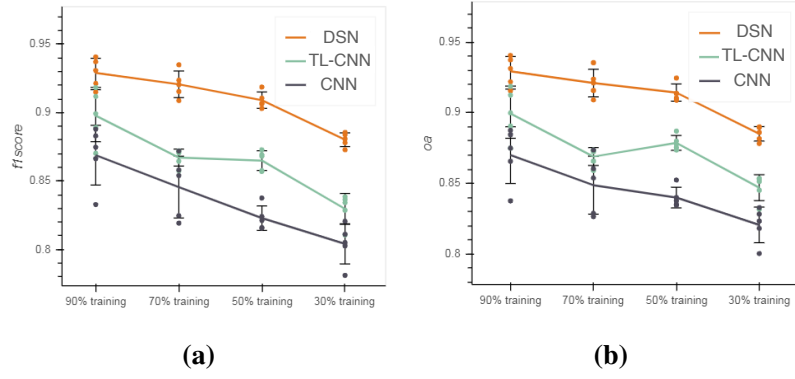


图 4.11 DSN 和 CNN 在 S1 数据集上的表现。(a) F1 分数, (b) 平均准确率。

Figure 4.11 The performance of DSN and CNN on S1 dataset. (a) F1-score, (b) Overall accuracy.

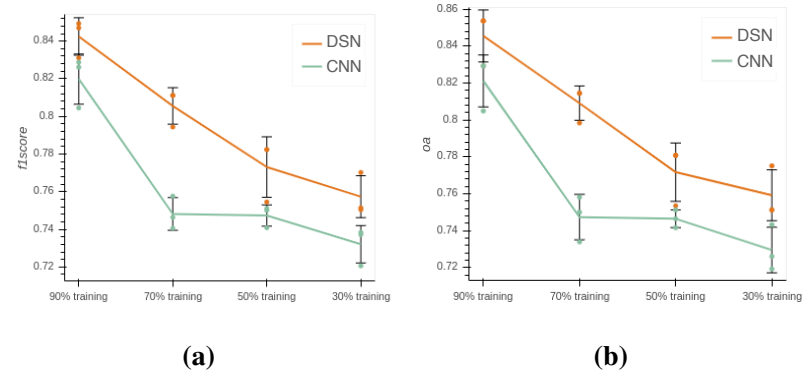


图 4.12 DSN 和 CNN 在 TSX 数据集上的表现。(a) F1 分数, (b) 平均准确率。

Figure 4.12 The performance of DSN and CNN on TSX dataset. (a) F1-score, (b) Overall accuracy.

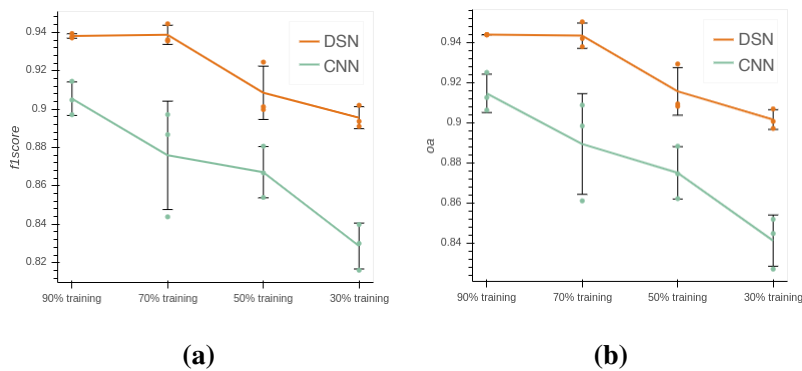


图 4.13 DSN 和 CNN 在 OPS 数据集上的表现。(a) F1 分数, (b) 平均准确率。

Figure 4.13 The performance of DSN and CNN on OPS dataset. (a) F1-score, (b) Overall accuracy.

现出明显的优势。TSX 和 OPS 数据集中的类别均为人造目标, 因此从图4.15(c)和图4.15(d)中也能观察到类似结果。只用少量数据(数据集的 30%)进行训练的时候, 由于神经网络过拟合, 只基于空间特征的普通 CNN 网络在人造目标上的泛化能力急剧降低, 而 DSN 借助频域中的信息将模型的性能提升, 在少量训练数据的情况下也能表现出对人造目标较好的区分能力。S1 数据集中各类别的准确率和 F1 分数记录在表4.4和表4.5中。

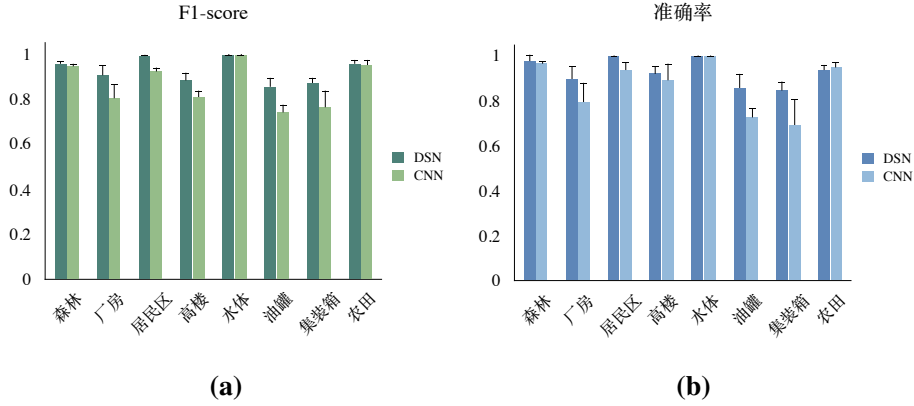


图 4.14 S1 数据集用 90% 的数据进行训练时, 每类地物在 DSN 和 CNN 模型中的准确率和 F1 分数。(a)F1 分数, (b) 准确率。

Figure 4.14 The accuracy and F1-score of DSN and CNN in each class, with 90% data of S1 dataset for training. (a) F1-score, (b) Accuracy.

4.4.4 基线模型对比

上节将 DSN 与基于幅度信息的 CNN 和 TL-CNN 进行了对比, 说明了 DSN 在人造目标类识别的优越性。本节比较了 DSN 和复数数据驱动的 CV-CNN, 以及只基于频域信号的 F-CNN 模型, 并对结果进行分析。

图4.16(a)展示了三个数据集在 DSN 和 CV-CNN 模型中的测试准确率, DSN 的表现明显优于 CV-CNN。DSN 和 CV-CNN 均采用了复数 SAR 数据来进行 SAR 图像分类任务的学习, 并具有类似的网络结构和参数体积。不同的是, CV-CNN 直接从复数数据中学习复网络参数, 是一种完全数据驱动的方法, 依靠输入数据对网络各层的权重进行学习, 得到从复数图像到类别空间的端到端映射, 在大规模数据可利用的情况下, 通过设计良好的网络结构, 选择合适的学习参数进行拟合。相比之下, DSN 框架涵盖了一些理论性的先验知识, 例如运用时频分析方法进行连续子带分解, 揭示 SAR 图像在不同子视条件下的散射变化。DSN 算法

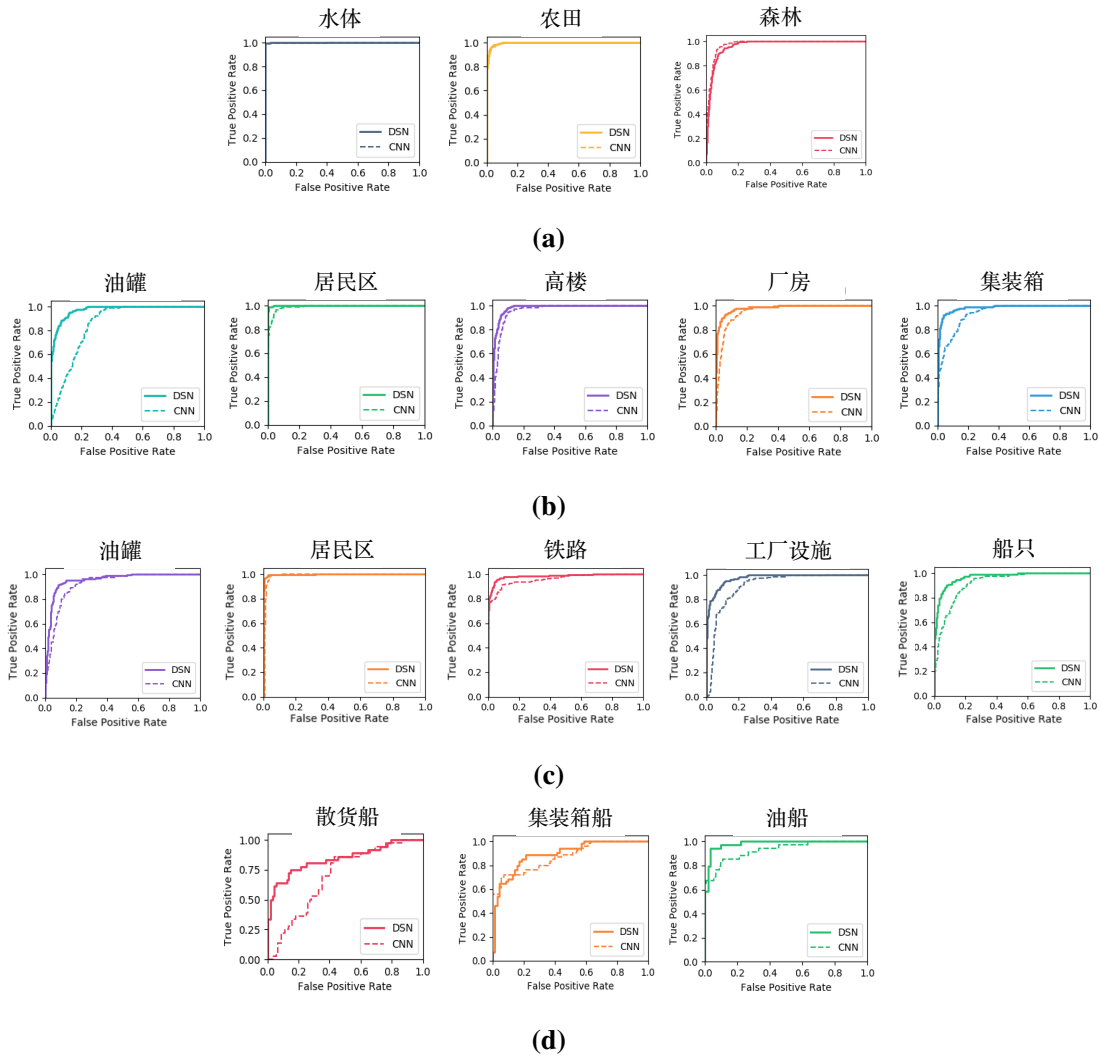


图 4.15 DSN 和 CNN 模型的 ROC 曲线。(a) S1 数据集自然地物类别，(b) S1 数据集人造目标类别，(c) TSX 数据集，(d) OPS 数据集。

Figure 4.15 ROC curve of DSN and CNN. (a) Natural classes of S1 dataset, (b) Man-made target classes of S1 dataset, (c) TSX dataset, (d) OPS dataset.

从复数 SAR 数据到地物类别空间经历了几个可解释的处理流程和学习阶段，与纯数据驱动的 CV-CNN 方法相比并不完全是一个“黑盒子”。DSN 的设计既考虑了 SAR 图像特殊的成像原理，运用了已知的物理模型，同时也将难以解释的过程（如特征提取）用小型的卷积神经网络来代替，采用数据驱动的方法，协同完成 SAR 图像分类任务。DSN 减少了从头训练 CV-CNN 深度网络的大数据需求，因此在带标签 SAR 图像较少的情况下能更好地克服过拟合问题，取得更优的结果。

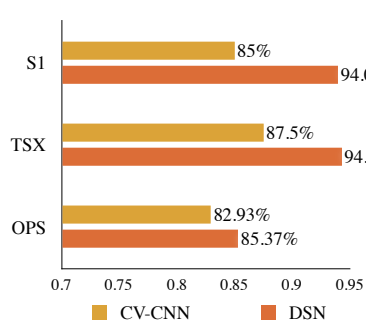
本小节还对比了只利用频域信号进行学习的 F-CNN、只利用空间幅度信息

表 4.4 S1 数据集在 DSN、TL-CNN 和 CNN 中各类别测试集的表现 (准确率%)

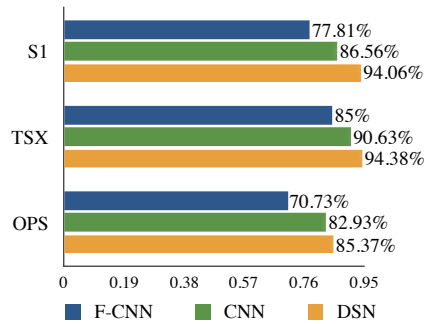
Table 4.4 The test accuracy of DSN, TL-CNN, and CNN on S1 dataset.

类别	90% 训练			70% 训练		
	DSN	TL-CNN	CNN	DSN	TL-CNN	CNN
森林	98±2.74	99.50±1.12	96.50±1.37	95.16±2.84	98.11±1.37	96.42±0.94
水体	100±0	100±0	100±0	99.58±0.58	99.79±0.47	99.79±0.47
农田	93.50±2.24	94.50±2.09	95.00±3.06	94.32±2.42	96.42±1.76	99.79±0.47
厂房	90±5.59	81±8.77	79.50±4.47	85.68±2.94	75.16±5.29	80.00±6.66
居民区	100±0	97.50±3.06	94.00±5.18	99.79±0.47	96.84±1.05	94.95±1.73
高楼	92.00±3.26	91.50±7.42	89.00±5.76	91.79±2.92	87.79±4.97	85.26±6.00
油罐	85.50±6.71	82.00±4.11	72.50±6.85	83.79±1.76	74.95±4.90	64.21±2.35
集装箱	84.50±4.11	73.50±10.98	69.50±9.25	85.05±7.64	66.11±5.78	64.00±17.25
平均值	92.94±1.05	89.94±1.75	87.00±2.01	92.21±0.94	86.89±0.64	84.87±2.05

类别	50% 训练			30% 训练		
	DSN	TL-CNN	CNN	DSN	TL-CNN	CNN
森林	94.57±1.92	96.67±2.33	96.42±2.74	80.70±4.07	95.68±2.21	91.98±1.77
水体	99.66±0.31	99.77±0.31	99.77±0.31	99.67±0.18	99.43±0.22	99.67±0.34
农田	92.32±1.77	94.92±2.31	91.35±3.20	90.27±2.55	91.12±3.31	90.04±2.96
厂房	86.67±3.27	78.74±3.75	74.59±4.52	86.03±3.05	73.85±1.15	74.87±2.73
居民区	99.77±0.31	96.93±1.03	96.25±1.78	99.67±0.18	94.72±2.97	95.85±2.12
高楼	89.03±2.57	88.39±5.80	81.94±4.48	87.85±2.63	82.92±2.20	72.85±4.41
油罐	81.64±4	70.41±6.75	61.52±2.56	78.66±2.90	69.29±2.49	64.69±5.28
集装箱	82.78±2.82	64.63±7.21	55.93±7.48	80.26±4.90	55.63±6.98	53.25±9.21
平均值	90.80±0.66	86.31±0.61	82.22±0.85	87.89±0.59	82.83±1.13	80.40±1.46



(a)



(b)

图 4.16 DSN 与基线模型的平均准确率对比。(a) CV-CNN 和 DSN, (b) CNN、F-CNN 和 DSN。

Figure 4.16 The comparison with baseline models. (a) CV-CNN and DSN, (b) CNN, F-CNN, and DSN.

表 4.5 S1 数据集在 DSN、TL-CNN 和 CNN 中各类别测试集的表现 (F1 分数)

Table 4.5 The F1-score of DSN, TL-CNN, and CNN on S1 dataset.

类别	90% 训练			70% 训练		
	DSN	TL-CNN	CNN	DSN	TL-CNN	CNN
森林	0.958±0.011	0.959±0.011	0.949±0.033	0.957±0.014	0.96±0.014	0.945±0.014
水体	1±0	1±0	1±0	0.998±0.003	0.997±0.003	0.998±0.003
农田	0.959±0.014	0.96±0.018	0.955±0.023	0.963±0.012	0.957±0.010	0.946±0.022
厂房	0.907±0.046	0.832±0.065	0.804±0.052	0.876±0.013	0.770±0.009	0.786±0.033
居民区	0.990±0.010	0.958±0.014	0.926±0.037	0.974±0.011	0.953±0.005	0.930±0.017
高楼	0.887±0.029	0.844±0.025	0.813±0.045	0.877±0.030	0.813±0.030	0.785±0.014
油罐	0.858±0.035	0.817±0.036	0.741±0.053	0.841±0.015	0.737±0.016	0.667±0.033
集装箱	0.873±0.023	0.811±0.075	0.762±0.0457	0.873±0.023	0.749±0.038	0.705±0.096
平均值	0.929±0.011	0.898±0.019	0.869±0.022	0.921±0.009	0.867±0.006	0.845±0.022
类别	50% 训练			30% 训练		
	DSN	TL-CNN	CNN	DSN	TL-CNN	CNN
森林	0.930±0.017	0.942±0.007	0.934±0.021	0.844±0.033	0.917±0.025	0.900±0.020
水体	0.998±0.001	0.995±0.002	0.996±0.001	0.997±0.001	0.995±0.001	0.995±0.002
农田	0.939±0.017	0.952±0.009	0.935±0.023	0.873±0.025	0.926±0.019	0.910±0.019
厂房	0.864±0.014	0.785±0.017	0.744±0.013	0.851±0.011	0.735±0.009	0.719±0.017
居民区	0.982±0.007	0.947±0.006	0.929±0.013	0.971±0.010	0.928±0.017	0.915±0.012
高楼	0.871±0.012	0.827±0.006	0.767±0.013	0.860±0.021	0.803±0.009	0.739±0.037
油罐	0.828±0.017	0.741±0.020	0.642±0.016	0.806±0.016	0.690±0.015	0.632±0.035
集装箱	0.860±0.012	0.726±0.054	0.632±0.058	0.838±0.025	0.640±0.045	0.618±0.059
平均值	0.909±0.006	0.865±0.007	0.823±0.009	0.880±0.005	0.829±0.011	0.804±0.014

进行学习的 CNN 和 DSN 的性能, 如图4.16(b) 所示。在三个数据集上, F-CNN 的效果都不如 CNN, 结合空间幅度特征和频域特征的 DSN 表现最好。在 SAR 图像分类任务中, 地物类别与图像的纹理密切相关, 传统方法凭借 SAR 图像的幅度纹理特征即能得到较好的分类结果。本章提出的频域信号利用 SAR 的特殊机制描述了一幅 SAR 图像的子带图像在距离向和方位向上的变化特性, 这种特性与目标的组成结构、材质等物理性质关联较大, 而与抽象的地物语义类别标签相差较远。F-CNN 忽略了合成孔径图像的空间特征而只关注频域信号, 在抽象语义层面对 SAR 图像的区分度并不高。但在密集的城市地区, 面对散射复杂的目标和场景 SAR 图像, CNN 仅凭空间纹理信息去区分难度较大, 加入频域信号予以帮助使得 DSN 能获取更多的知识, 辅助空间信息对 SAR 图像的类别进行判

别。

4.4.5 可视化分析

本节将进一步对实验结果进行可视化分析和讨论。首先将 DSN 和 CNN 算法中 G 部分最后一个残差瓶颈模块的输出特征进行可视化，S1 测试集的可视化结果如图4.17所示，红色和黄色分别表示森林和农田，可以看到在 CNN 模型中二者除了在各自的特征簇中混杂了少许的样本之外，主要的聚簇能够清晰地被分开。而在 DSN 模型中，森林和农田两个特征簇的边界重合，区分度稍低。图4.15和表4.4也反映出 DSN 在自然地物类别（例如水体、农田、森林）与传统的 CNN 算法相比并无优势。由于多数自然地物在不同子带下不具有高相关性，表现为平稳目标，其频域信号区分较小。图4.18展示了两个森林和农田样本在相同位置上的频域信号幅度图，两种自然地物在频谱上会表现出相似的特征，由此推出自然地物在频域上的特征不足以给具体的地物类别分类提供足够的额外信息。

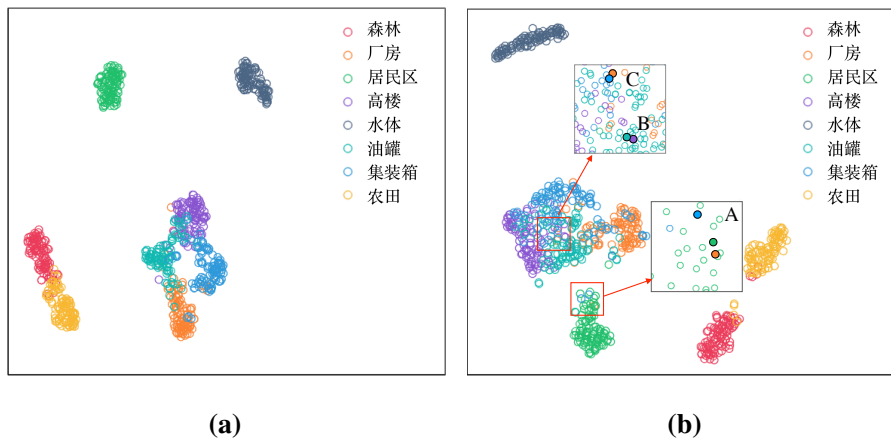


图 4.17 S1 数据集在 DSN 和 CNN 中的特征分布。(a) DSN, (b) CNN。

Figure 4.17 The feature visualization of S1 dataset in DSN and CNN (a) DSN, (b) CNN.

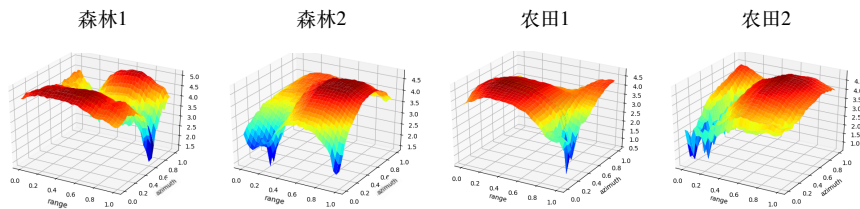


图 4.18 自然地物样本频域信号幅度图可视化

Figure 4.18 The frequency domain signal visualization of natural samples.

而对于另外五种人造目标相关的地物类别：居民区、集装箱、高楼、厂房

以及油罐，频域信号起到了很好的辅助作用。从图4.17(a)中可以看到，在DSN模型中居民区能够非常明显地与其他几类人造目标分开，而图4.17(b)中显示的CNN模型下居民区的特征簇和其他四类人造目标相距较近，其中还混淆了一些集装箱和厂房的样本，如情况A标注的几个样本所示。集装箱、高楼、厂房及油罐这几类人造目标具有极为复杂的散射表现，在幅度图像中表现出较难理解的纹理特征，因而CNN在识别这几个类别时产生了较大的困难，图4.17(b)中显示这四类的特征簇几乎都混在了一起难以区分，有的不同类别的样本却具有非常相似的空间特征，如情况B和情况C标注的几个样本所示。与之形成对比的是，图4.17(a)显示的DSN模型中，这四类样本的特征簇聚类更紧密，也更容易区分。

图4.19给出了情况A、B和C中标记的几个样本，以及DSN和CNN对它们的预测结果，红圈标记了图中容易被CNN混淆的纹理，同时附上红圈标注位置对应的频域信号。情况A如图4.19(a)所示，厂房和集装箱样本都被CNN以78.16%和71.77%的概率预测为居民区，DSN则分别以75.47%和76.06%的概率做出了正确的预测。图中同时还展示了和这两个样本特征相似的居民区样本，在CNN和DSN中均得到了正确的预测结果。通过与光学图像的比对，红圈标记的矩形特征在居民区样本中，对应街区的大致形状，矩形中较亮的散射点反映出分布其中的房屋；在厂房样本中，矩形特征对应厂房屋顶，厂房的一侧边缘因为多次散射表现出一条亮线；在集装箱样本中，矩形四周边缘因为周围密集排列的集装箱产生多次散射而表现出较强的散射。情况B中的油罐和高楼SAR图像中均出现了强散射点以及十字星，误导CNN做出错误的判断，将高楼也判别为了油罐。情况C类似，DSN以99.77%的高预测概率正确预测了集装箱样本，而CNN则只依赖空间信息给出了一个概率为67.25%的预测值，而且在与之类似的厂房样本的预测中得到了错误的结果。

几个SAR切片的强度图像具有相似的空间特征，其对应位置上的频域信号却表现出不同的散射特性。居民区的各类目标具有各向同性的散射性质，包括墙体与粗糙地面所形成的二面角以及圆柱体的路灯等，频域信号中显示出对应位置在各子带中的变化较为平缓，情况B中独立分布的油罐与之类似。大型厂房的方位和屋顶极为规则的结构特征、集装箱的金属材料以及密集的排列、高楼集中区域的多径散射，都会导致各向异性或非平稳性，相比于居民区中的目标散射特性更复杂，在频域信号中有所体现。因此基于空间和频域信息的DSN在这一

点上具有一定的优势。

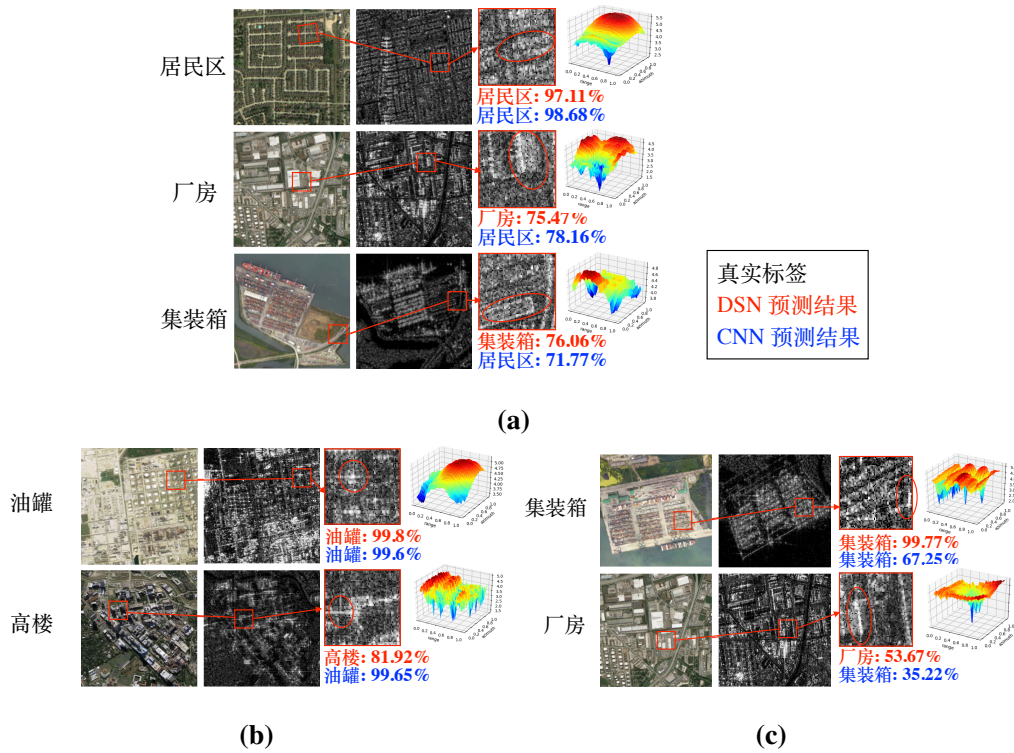


图 4.19 图 4.17(b) 中所示 A、B、C 样本在 DSN 和 CNN 中的预测结果对比及频域信号分析。

(a) 示例 A, (b) 示例 B, (c) 示例 C。

Figure 4.19 The elaborate explanation of case A, B, and C in Fig. 4.17(b). (a) The case A, (b) The case B, (c) The case C.

类似的, 图4.20展示了 TSX 数据集在 DSN 和 CNN 情况下的特征分布。TSX 数据集中的五种类别都与人造目标有关, 其中由于一些工业设施比较靠近油罐的区域, 在 CNN 的学习过程中容易将二者混淆, 如图4.20(b) 灰色和红色所示; 另一方面由于一些工业设施中大量密集交叉的金属架与有的船只上的金属部件类似, 都呈现出杂乱的强散射, 因此在视觉上难以分清, 如图4.20(b) 中红色和紫色所示。而图4.20(a) 中给出的 DSN 的特征可视化分布结果, 可以看到工业设施特征的聚簇更加紧凑, 并且和油罐以及船只的区分度得到了提升。

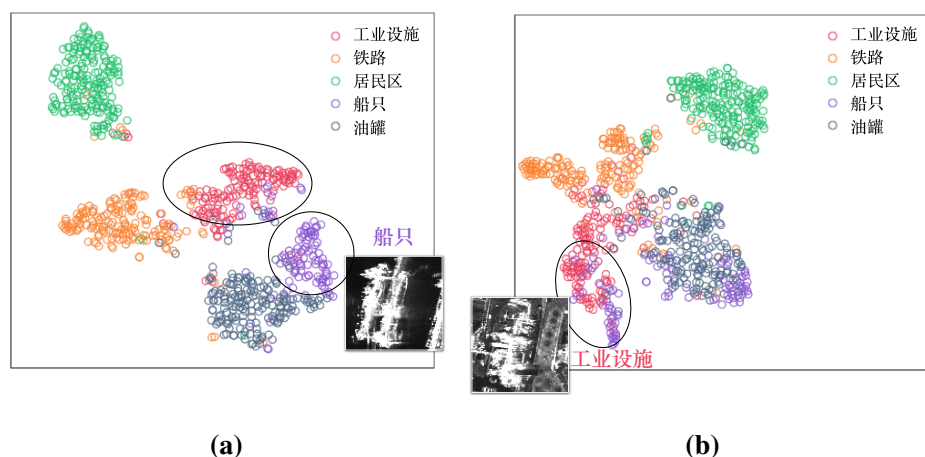


图 4.20 TSX 数据集在 DSN 和 CNN 中的特征分布。(a) DSN, (b) CNN。

Figure 4.20 The feature visualization of TSX dataset in DSN and CNN (a) DSN, (b) CNN.

4.5 本章小结

针对 SAR 图像目视解译困难，在深度学习应用中信息利用不充分的问题，本章提出了一个新型的深度学习框架 Deep SAR-Net (DSN)，对复数 SAR 图像进行二维连续子带分解，联合合成孔径全分辨率 SAR 图像的幅度信息和频域的子带信号，分别设计四个子网络进行学习，完成 SAR 图像分类任务。该框架对比传统的 CNN 方法，在人造目标的区分上带来了明显的提升。和纯复数数据驱动的复卷积网络相比，融合 SAR 特殊的成像机制和时频分析理论的 DSN 降低了对大数据的要求，在标记困难的 SAR 图像分类任务中具有实际应用价值。

本章的主要工作已成文发表，并在国际会议上进行口头报告：

- HUANG Z, DATCU M, PAN Z, LEI B. Deep SAR-Net: Learning objects from signals [J], **ISPRS Journal of Photogrammetry and Remote Sensing**, 2020, 161:179-193
- HUANG Z, DUMITRU C O, PAN Z, LEI B, DATCU M. A novel deep learning framework based on transfer learning and joint time-frequency analysis [C], **TerraSAR-X/TanDEM-X Science Team Meeting**, 2019

第5章 子带散射特性提取

5.1 引言

与光学遥感图像不同, SAR 作为主动微波成像系统具有特殊的成像模式, 其成像结果经过处理得到的 SAR 图像反映了地表及地面目标对电磁波的后向散射, 充分理解 SAR 图像反映的物理散射特性对 SAR 图像解译具有重要意义。对极化 SAR 图像, 基于极化物理模型的极化分解理论和极化散射特性分类方法 (Cloude 和 Pottier, 1997; Lee 等, 2004; Yamaguchi 等, 2006; Chen 和 Sato, 2012) 揭示了 SAR 图像中的物理散射特性, 并广泛应用在植被鉴别 (Sato 等, 2009)、建筑物密集区域提取 (Ratha 等, 2019)、植被中的人造目标辨别 (Guinvarc'h 和 Thirion-Lefevre, 2017) 等任务中。同时也有研究结合了极化 SAR 图像的极化散射特性和空间信息来进行极化 SAR 图像的分类 (Zhou 等, 2016b; Du 等, 2015)。

近年来, 多颗高分辨率的 SAR 卫星相继升空, 例如 TerraSAR-X (TerraSAR-X Science Service System) 和高分三号卫星 (Sun 等, 2017) 等, 在某些模式下能提供超高分辨率 (Very High Resolution, VHR) 的 SAR 图像, 可以对地面目标进行细节丰富的成像, 例如 TerraSAR-X 聚束模式分辨率能达到亚米级。然而很多情况下 VHR 和多种极化方式难以同时满足, 在面向 VHR 的单极化 SAR 图像解译开展的基于深度学习方法的研究中, 考虑到超高分辨率带来的丰富的纹理信息, 多数工作针对 SAR 图像的幅度信息对空间纹理特征进行学习 (Huang 等, 2017b; An 等, 2018)。在这种思路下, SAR 图像被当做传统的“图像”来进行处理, 忽略了 SAR 图像的物理散射特性。

上一章的研究提出, 利用 SAR 图像的二维连续子带分解得到的高维时频信号能有效辅助传统的基于幅度信息的 CNN 网络进行 SAR 图像分类。这意味着复数 SAR 图像的时频分析理论揭示了一些图像域看不到的信息, 相关研究也表明利用时频分析方法能够描述目标在复数 SAR 图像中对不同方位视角以及不同频率的响应变化 (Tupin 和 Tison, 2004)。另外, Spigai 等 (2011) 根据复数 SAR 图像的时频分析得到的雷达频谱图提出了四种典型目标及对应的雷达频谱图模式, 并对实际 SAR 场景中的目标进行了分析, 讨论了其与极化散射特性分析结果的关联性, 但这篇论文仅局限于几种典型目标, 无法像极化散射特性分类方法一

样，对实际 SAR 场景进行逐目标的分析。

本章在上一章的研究驱动下，希望更深层次地挖掘复数 SAR 图像的连续二维子带分解过程能够提供什么样的散射信息，在缺少物理模型支撑的情况下，用数据驱动的深度学习方法发现高分辨率单极化 SAR 图像潜在的物理散射特性。本章提出基于时频分析理论的无监督多层次深度嵌入聚类方法（hierarchical deep embedding clustering based on time-frequency analysis, HDEC-TFA），首先利用时频分析理论对目标提取子带散射图，体现目标在不同方位视角和不同电磁波频段下后向散射强度的变化情况，然后采取两个阶段的深度嵌入聚类方法同时学习子带散射图的特征和聚类中心，最后将结果可视化，形成子带散射特性分类结果图。为了评估所提方法有效地学习到了具有实际意义的物理散射特性，本章还设计了一种基于信息论的定量化评估方法，将极化分析的结果作为参考进行讨论。

本章内容安排如下，首先在第二节简单介绍了极化 SAR 图像的散射特性分类方法，用于生成参考数据，第三节详细介绍了所提的子带散射特性无监督分类方法，第四节基于高分三号的 SAR 图像数据进行了实验，给出实验结果和分析讨论的详细内容，最后第五节为本章小结。

5.2 极化 SAR 图像的散射特性分类

本节简单介绍极化 SAR 的基本概念和极化 SAR 图像散射特性分类的典型方法，为后续实验提供极化散射特性参考。

5.2.1 极化 SAR 的基本概念

极化合成孔径雷达（Polarimetric SAR, PolSAR）发射多种极化方式的电磁波，电磁波与地面目标相互作用后产生不同极化状态的回波，由此获得同时包含图像功率和目标物理特性的复信号。在全极化 SAR 中，一般采用 2×2 的 Sinclair 极化后向散射矩阵来描述单个像素点的散射特性，包含目标的全部极化信息，记为

$$\mathbf{S} = \begin{bmatrix} S_{HH} & S_{HV} \\ S_{VH} & S_{VV} \end{bmatrix} \quad \dots (5.1)$$

其中 S_{HV} 表示水平极化通道（H）接收的垂直极化通道（V）发射回波的复散射系数， S_{HH} 表示水平极化通道（H）接收的水平极化通道（H）发射回波的复散射系数，其它以此类推。散射矩阵 \mathbf{S} 满足互易性，即 $S_{HV} = S_{VH}$ 。已知了散射矩

阵 S ，极化 SAR 图像中的分布式目标就可以用多视的极化协方差矩阵 C 、极化相干矩阵 T 和 Kennaugh 矩阵 K 来描述。

对散射矩阵 S 矢量化可以得到极化测量矢量

$$\vec{\Omega} = \begin{bmatrix} S_{HH} & \sqrt{2}S_{HV} & S_{VV} \end{bmatrix}^T. \quad \dots (5.2)$$

则目标的极化协方差矩阵 C 表示为矢量 $\vec{\Omega}$ 与自身共轭转置的外积，记为

$$C = \langle \vec{\Omega} \cdot \vec{\Omega}^\dagger \rangle, \quad \dots (5.3)$$

其中 $\langle \cdot \rangle$ 表示空间平均， \dagger 表示复共轭转置。

目标的极化相干矩阵 T 可以由 Pauli 基得到，Pauli 基矢量表示为

$$\vec{k} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} & S_{HH} - S_{VV} & 2S_{HV} \end{bmatrix}^T, \quad \dots (5.4)$$

则极化相干矩阵可以写为

$$T = \langle \vec{k} \cdot \vec{k}^\dagger \rangle. \quad \dots (5.5)$$

定义 4×4 的 Kennaugh 矩阵 K 为

$$K = \frac{1}{2} A^* (S \otimes S^*) A^\dagger, \quad \dots (5.6)$$

其中

$$A = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & j & -j & 0 \end{bmatrix}, \quad \dots (5.7)$$

\otimes 表示 Kronecker 积， $*$ 表示复共轭， $j = \sqrt{-1}$ 。Kennaugh 矩阵为实矩阵，可以由极化相干矩阵 T 得到：

$$K = \begin{bmatrix} \frac{T_{11}+T_{22}+T_{33}}{2} & \Re(T_{12}) & \Re(T_{13}) & \Im(T_{23}) \\ \Re(T_{12}) & \frac{T_{11}+T_{22}-T_{33}}{2} & \Re(T_{23}) & \Im(T_{13}) \\ \Re(T_{13}) & \Re(T_{23}) & \frac{T_{11}-T_{22}+T_{33}}{2} & -\Im(T_{12}) \\ \Im(T_{23}) & \Im(T_{13}) & -\Im(T_{12}) & \frac{-T_{11}+T_{22}+T_{33}}{2} \end{bmatrix} \quad \dots (5.8)$$

其中 \Re 和 \Im 分别表示取实部和虚部。

基于散射特性和极化系统物理建模的极化分析可以应用到所有的极化 SAR 数据中，与传感器或数据的特殊性无关。很多极化分析的研究利用极化协方差矩阵或极化相干矩阵的分解来揭示物理散射特性，例如 $H/A/\alpha$ 分解 (Cloude 和 Pottier, 1997)，Freeman-Durden 分解 (Freeman 和 Durden, 1998) 和 Yamaguchi 等 (2006) 提出的四分量分解等。Lee 等 (2004, 1999) 在极化分解的基础上提出著名的基于 Wishart 分布的无监督分类算法，广泛应用于极化 SAR 图像的散射特性分类。

地面实际目标由于复杂的结构和散射特性，在极化信息中存在非单一的散射响应。一般可以将复杂的散射用几种基本散射机制的组合来表示，例如表面散射 (Surface Scattering)、二面角散射 (Dihedral Scattering)、体散射 (Volume Scattering)、螺旋体散射 (Helix Scattering) 等。You 等 (2014) 将 Kennaugh 矩阵分解为两个部分，即 $\mathbf{K} = \mathbf{K}^0 + \Delta\mathbf{K}$ 。其中 \mathbf{K}^0 代表与 \mathbf{K} 最近的相干目标， $\Delta\mathbf{K}$ 看做一个噪声矩阵。分解之后的相干目标 \mathbf{K}^0 具有两个特殊性质，其总功率 Span 与实际目标 \mathbf{K} 的总功率相同，且相干目标 \mathbf{K}^0 代表的散射特性与实际目标 \mathbf{K} 的主要散射特性一致。

5.2.2 极化散射特性的无监督分类

本节简要介绍 Ratha 等 (2018) 提出的基于 Kennaugh 矩阵和测地距离 (Geodesic Distance) 的 Wishart 无监督分类方法，在实验部分将用此方法生成全极化 SAR 图像的散射特性图，作为参考数据开展分析和讨论。

定义三个 Kennaugh 矩阵用来描述典型的散射机制，例如可用三面角反射器建模的奇次散射 (Odd-bouncing, 记为 O) 对应的 \mathbf{K} 矩阵定义为：

$$\mathbf{K}_O = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix}, \quad \dots (5.9)$$

可用二面角反射器建模的二次散射 (Double-bouncing, 记为 D) 对应的 \mathbf{K} 矩阵定

义为

$$\mathbf{K}_D = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad \dots (5.10)$$

可用大量均匀的分布式随机朝向偶极子散射器建模的体散射（Volume-bouncing, 记为 V）对应的 \mathbf{K} 矩阵定义为

$$\mathbf{K}_V = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \quad \dots (5.11)$$

任意两个矩阵之间的测地距离定义为

$$\text{GD}(\mathbf{K}_1, \mathbf{K}_2) = \frac{2}{\pi} \cos^{-1} \left(\frac{\text{Tr}(\mathbf{K}_1^T \mathbf{K}_2)}{\sqrt{\text{Tr}(\mathbf{K}_1^T \mathbf{K}_1)} \sqrt{\text{Tr}(\mathbf{K}_2^T \mathbf{K}_2)}} \right). \quad \dots (5.12)$$

由此，计算待观测的 \mathbf{K} 矩阵与典型目标（ \mathbf{K}_O , \mathbf{K}_D , 和 \mathbf{K}_V ）之间的归一化测地距离，并用总功率 Span 调制，记为 w_i ，其中 $i \in \{O, D, V\}$ ：

$$w_i = \text{SPAN} \times \frac{1 - \text{GD}(\mathbf{K}, \mathbf{K}_i)}{\sum_i (1 - \text{GD}(\mathbf{K}, \mathbf{K}_i))}, \quad \dots (5.13)$$

总功率 SPAN 的定义为

$$\text{SPAN} = |S_{HH}|^2 + |S_{HV}|^2 + |S_{VH}|^2 + |S_{VV}|^2. \quad \dots (5.14)$$

对每个像素点计算 \mathbf{K} 矩阵以及归一化测地距离 w_i ，根据 w_i 的最大值对应的 i 选择最主要的散射类别 $i \in \{O, D, V\}$ 作为像素点的初始类别。与 Ratha 等 (2018) 所提方法不同的是，如若 w_i 的值满足 $\frac{\max\{w_O, w_D, w_V\}}{w_O + w_D + w_V} \leq C$ (这里取 $C = 0.5$)，本节将其定义为混合类（Mix-bouncing, 记为 M）。初始化散射类别之后，将每个类别分别进行重聚类（Clustering）和簇合并（Cluster Merging），然后应用 Wishart 分类方法进行迭代，得到最终的分类结果。

5.3 子带散射特性聚类方法

5.3.1 流程总览

本章提出针对单极化 SAR 图像的子带散射特性聚类方法，基于 SAR 复数图像的时频分析理论提取每个像素对应的子带散射图（sub-band scattering pattern），利用深度神经网络自动学习子带散射图的特性并进行多层级的聚类，得到子带散射特性分类结果图。为了与极化 SAR 图像的散射特性分类结果进行对比，本章选取全极化 SAR 图像 HH 通道的数据作为单极化 SAR 图像进行实验，设计量化评估方法与全极化 SAR 图像的极化散射特性分类结果进行对比和分析。

本方法的主要流程如图5.1所示，以下对所提方法的工作流程做简要介绍。

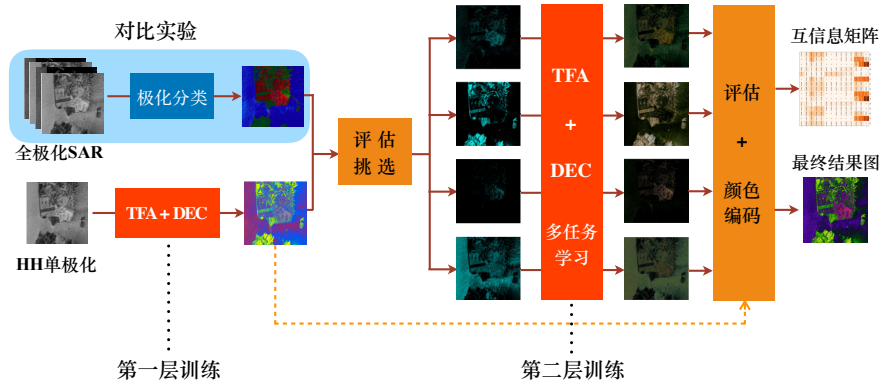


图 5.1 基于时频分析的无监督子带散射特性聚类方法主要流程图

Figure 5.1 The workflow overview of the proposed sub-band scattering clustering method.

1) 全极化 SAR 图像的极化散射特性分类图提取

对 HH/HV/VH/VV 四通道的全极化 SAR 图像，采用第5.2.2小节介绍的极化分析方法 (Ratha 等, 2018) 获得散射特性分类图，共包括奇次散射 (O)、二次散射 (D)、体散射 (V) 和混合散射 (M) 四种类别；

2) 单极化 SAR 图像的子带散射特性提取

对 HH 通道的单极化 SAR 图像，基于二维时频分析理论提取每个像素点在距离向和方位向两个维度的不同频带中的散射表现，生成子带散射图；

3) 第一阶段深度嵌入聚类网络训练

设计深度卷积神经网络用于子带散射图的特征提取和无监督自动分类，采用深度嵌入聚类方法 (Deep Embedding Clustering, DEC) 进行训练，同时学习特征表示和特征聚类中心，得到第一阶段的散射特性分类图；

4) 评估和挑选

将第一阶段的散射特性分类图与全极化 SAR 图像的散射特性分类结果进行对比，量化评估生成互信息矩阵，并挑选合适的类别进行第二阶段的训练；

5) 第二阶段深度嵌入聚类网络训练

对 4) 中挑选出的类别以多任务学习 (Multi-task learning) 的方式重复步骤 3) 的训练，得到每类下的不同子类；

6) 评估和颜色编码

与全极化 SAR 图像的散射特性分类结果进行对比，量化评估生成互信息矩阵；将分类结果进行颜色编码，生成最终的子带散射特性分类图。

5.3.2 目标的子带散射图

本节介绍 SAR 图像中目标的子带散射图提取方法，利用目标在距离和方位向两个维度上不同频段的后向散射变化来描述目标的物理特性。

第4.2节介绍了基于时频分析方法的 SAR 图像二维连续子带分解，对一张合成孔径的复数图像 $C(x, y)$ 通过短时傅里叶变换进行频域滤波，得到一系列二维连续频带的子带图像。目标的子带散射图提取与之类似，目的是针对 SAR 图像中的某目标，即像素点 (x_0, y_0) ，提取其在距离和方位向每个二维子带的后向散射强度变化。具体实现如图5.2所示。

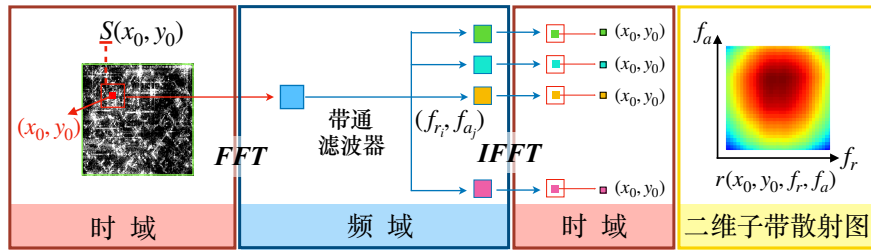


图 5.2 提取目标子带散射图的时频分析方法

Figure 5.2 The time-frequency analysis method for extracting sub-band scattering pattern of target.

对于目标 (x_0, y_0) ，以此为中心在复数 SAR 图像上选取邻域 $s(x_0, y_0)$ ，通过傅里叶变换操作映射到频域得到

$$\mathcal{S}(f_r, f_a) = \text{FFT}[s(x_0, y_0)]. \quad \dots (5.15)$$

在频域用中心频率对在距离和方位向上连续变化的一系列带通滤波器 $w(f_r^i, f_a^j)$

进行二维滤波,得到滤波后的频域信号后再反变换到时域上。以上过程与第4.2节描述的二维连续子带分解一致,不同的是在这里取出每个子带图像在中心点 (x_0, y_0) 的散射强度,得到目标 (x_0, y_0) 在二维连续子带的散射图

$$r(x_0, y_0, f_r, f_a) = |\text{FFT}^{-1}[\mathcal{S} \cdot w(f_r, f_a)]|(x_0, y_0) \quad \dots (5.16)$$

如图5.2所示。

SAR 在距离向和方位向上的分辨率分别与发射线性调频信号的带宽和方位向波束宽度有关,因此高分辨率的 SAR 一般具有较大的距离向带宽和较宽的方位向波束。在合成的全分辨率 SAR 图像中,假设了目标具有平稳性特征,忽略了在变化的方位向视角和不同频率的发射电磁波下目标的散射变化。采用上述二维频率域分解的方法得到的二维连续子带散射图,揭示了目标在不同的方位向视角下对不同频率电磁波的散射强度响应,子带散射图体现出的在距离向和方位向上的散射强度变化反映了目标的物理特性,在这里将其称之为子带散射特性。

Spigai 等 (2011) 提出雷达频谱图 (radar spectrogram, 即本章提出的子带散射图) 具有四种基础的散射模式,对应不同的典型目标:

- **频率不变性 (frequency-invariant)** 子带散射图在距离向和方位向均表现出较稳定的特征,散射强度随频率变化起伏不大,一般对应具有各向同性的目标,如图5.3中目标 A——路灯所示;

- **二维多样性 (2D-variant)** 子带散射图在距离向和方位向上表现不稳定,散射强度随频率随机变化,无明显特征,一般对应 SAR 图像中的杂波信号,如图5.3中目标 B——裸露地面所示;

- **方位向多样性 (azimuth-variant)** 子带散射图在距离向上表现较稳定,散射强度随方位向频率变化较多,表明目标的散射表现容易受目标与 SAR 传感器相对位置的变化影响,多发生在一些各向异性的人造目标,或具有复杂结构的散射体上。图5.3中给出了一个船的例子,见目标 C。

- **距离向多样性 (range-variant)** 子带散射图在距离向上表现不稳定,目标对发射信号的频率响应敏感,一般对应具有共振结构的散射体,例如谐振球体、具有非常规则几何形态的屋顶等,如图5.3的目标 D——长方体的工业厂房屋顶。

在实际 SAR 图像中,除了这四种典型目标反映出来的散射模式,还存在其他更复杂的子带散射特性。特别是在密集的城市区域,紧密交错的非典型人造

目标使子带散射图表现出典型目标无法描述的特性，例如图5.3中的目标 E 和 F 所示。因此，本章拟采用数据驱动的非监督深度学习方法自动发现子带散射图中潜在的子带散射特性并进行分类，提出基于时频分析的多层级深度嵌入聚类方法（hierarchical deep embedding clustering based on time-frequency analysis, HDEC-TFA）。

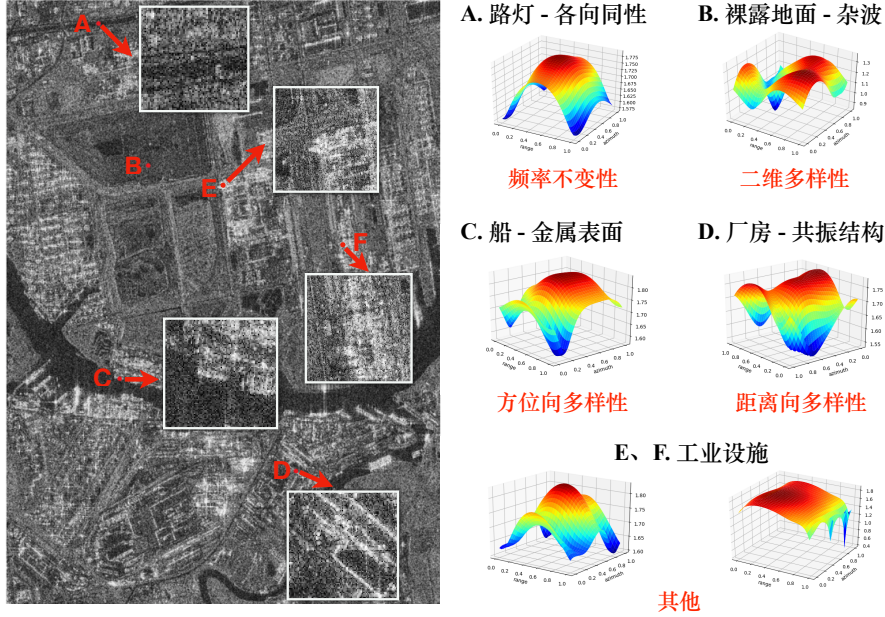


图 5.3 几种典型目标和非典型目标的子带散射图

Figure 5.3 The sub-band scattering patterns of some canonical and non-canonical targets.

5.3.3 多层级深度嵌入聚类

根据上一节的介绍，基于时频分析理论可以得到每个像素点对应的子带散射图，由此一幅复数 SAR 图像即能产生大规模的数据推动深度神经网络的训练。和极化分析类似，对一张单极化 SAR 图像也希望能对其中的所有目标进行散射特性的分类，不同的是极化分析依靠极化散射模型开展研究，具有坚实的物理模型理论基础，而本章提出的子带散射图缺乏足够的专家知识和先验信息，因此本节介绍一种无监督的深度学习方法对大规模的子带散射图进行自动的特征学习和分类。

深度嵌入聚类方法（Deep Embedding Clustering, DEC）由Xie 等 (2016) 提出，结合自编码网络和 K-means 算法，能够同时学习深度神经网络中的嵌入式特征和聚类算法中的聚类中心，如图5.4所示。本节基于此进行改进，提出面向子带散射图的多层级深度嵌入聚类方法。

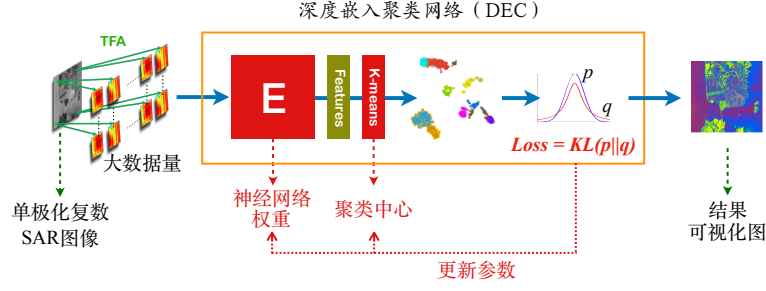


图 5.4 深度卷积嵌入聚类方法示意图

Figure 5.4 The deep convolutional embedding clustering method.

首先设计一个去噪堆栈卷积自编码器（Denoising stacked convolutional auto-encoder, DSCAE）来训练将子带散射图通过非线性映射，嵌入到低维空间的特征，如图5.5所示。考虑到子带散射图表征的散射特征远比普通自然图像或 SAR 图像的纹理简单，因此用于子带散射图特征提取的网络设计和训练方法均沿袭了第4.3.3节的内容。将 DSCAE 的编码部分记为 $\mathcal{G}_F(\theta)$ ，将子带散射图映射到 n 维的嵌入式空间得到嵌入特征 $\phi = \mathcal{G}_F(r) \in \mathbb{R}^n$ ，其中 θ 即为非线性映射的参数。然后利用 k -means 算法 (Hartigan 和 Wong, 1979) 对 n 维的嵌入特征进行聚类，得到 k 个初始化的聚类中心 $\{\mu_j\}_{j=1}^k$ 。由此获得了 θ 和 μ 两个参数的初始估计值。

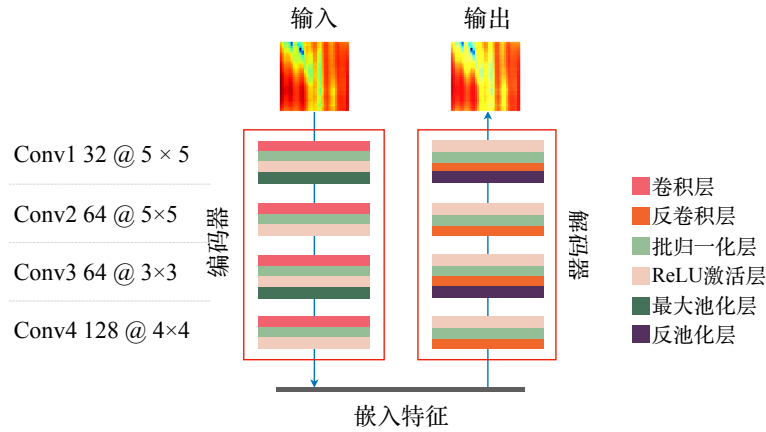


图 5.5 卷积自编码网络示意图

Figure 5.5 The convolutional auto-encoder network.

接下来对 θ 和 μ 进行参数优化。对于任意一个子带散射图的输入 r_i ，得到嵌入低维特征 ϕ_i 和被分配到的距离最近的聚类中心 μ_j 。用自由度为 1 的学生 t 分布作为核，来度量嵌入特征 ϕ_i 和聚类中心 μ_j 的相似度，表示将嵌入特征 ϕ_i 归

为第 j 个聚类中心 μ_j 的概率, 记为软分配 q_{ij}

$$q_{ij} = \frac{(1 + \|\phi_i - \mu_j\|^2)^{-1}}{\sum_{j'} (1 + \|\phi_i - \mu_{j'}\|^2)^{-1}}. \quad \dots (5.17)$$

优化时, 希望模型从高置信度的分配中进行学习, 因此定义一个目标辅助分布 P , 与软分配 Q 进行匹配:

$$p_{ij} = \frac{q_{ij}^2 / \sum_i q_{ij}}{\sum_{j'} (q_{ij'}^2 / \sum_i q_{ij'})}. \quad \dots (5.18)$$

然后将软分配和目标辅助函数之间的 KL 散度作为优化的目标函数, 即

$$Loss_{kl} = KL(P||Q) = \sum_i \sum_j p_{ij} \log \frac{p_{ij}}{q_{ij}}. \quad \dots (5.19)$$

在训练过程中, 非线性映射的参数 θ 和聚类中心 μ 根据目标函数 $Loss_{kl}$ 分别对 μ 和 ϕ 的求导, 通过动量随机梯度下降法进行学习。

在测试阶段, 对每个输入的子带散射图 r_i , 将其在 SAR 图像上对应的位置 (x_i, y_i) 赋予聚类的类别 j , 得到最终子带散射分类结果图。类比极化分析时对极化散射矩阵采取的多视平均操作, 测试时也将输入的子带散射图 r_i 以 3×3 的窗口做平均, 以得到更平滑的结果。

实验过程中发现, 很多与自然地物相关的反映二维杂波子带散射图在距离向和方位向上表现出的变化性不具有典型的特性, 形成较为明显的特征多样化、离散化。在这里将具有这种性质的子带散射类别称为“非平稳类”。当聚类中心个数 k 逐渐增大时, 这些子带散射图的类型会分裂成更多的小簇, 却阻止了“平稳类”的子带散射图进行更细致的分析。实验还发现, 这些“平稳类”的子带散射图与极化分析结果中的某些类别具有更大的关联性, 值得更深层的研究。相关的实验在第5.4节中会有详细介绍。因此本方法决定采用多层级的训练模式, 对“平稳类”的子带散射图实行第二层的深度嵌入式聚类, 方便与极化散射特性进行更细致的对比。

对于多层级的训练模式, 在结束第一层训练之后将挑选合适的几种散射类别进入第二层的训练, 挑选原则在下一节中介绍。在第二层训练过程中, 采用多任务学习的方式同时进行, 即非线性映射网络的前两层参数共享, 新的聚类中心分别计算并更新。像这样平行的多任务学习方式能保证训练数量较少的类别时, 在非线性映射的特征学习部分避免过拟合, 增强模型的泛化能力。

5.3.4 互信息评估方法

前两节介绍的基于时频分析理论的多层级深度聚类方法创造性地提出将子带散射特性进行无监督分类，得到反映单极化 SAR 图像物理特性的散射类别。不同于极化分析方法基于发展较成熟的极化理论和物理模型，具有可解释性，本章提出的 HDEC-TFA 方法得到的散射类型是数据驱动的、基于学习的结果，目前无法用已有的先验知识去评估每种类别的准确性。虽然 Spigai 等 (2011) 提出了四种典型目标和对应的散射模式，但在真实的 SAR 场景中还包括其他更复杂的类别，而且对一个大的 SAR 场景逐目标进行散射模式的标注也是不可能的。另一方面，Spigai 等 (2011) 也指出，基于时频分析理论得到的典型散射模式与极化分析中的某些类别具有一定的联系。

因此本节提出用第5.2.2节介绍的全极化 SAR 图像的散射特性分类方法（这里记为 GD-Wishart 法）得到的极化散射特性图作为一个参考性的结果，设计一种基于信息论的评估方法对单极化 SAR 图像的子带散射特性提取结果进行评价。需要注意的是，这里并没有将极化 SAR 的散射特性作为真值，因为两种方法在本质上是不同的。本节提出的量化评估方法只是为了说明两种方法得到的结果具有一定的相关性，从而证实面向单极化 SAR 图像的 HDEC-TFA 方法能够学习到有切实物理意义的散射特性。

Valverde-Albacete 和 Peláez-Moreno (2019) 提出将监督分类的结果用基于信息论的定量描述方法进行评估，把从真实标签到预测标签信息量的传播效率用来衡量分类器的优劣。受 Valverde-Albacete 和 Peláez-Moreno (2019) 工作的启发，本节提出以下互信息评估方法。

首先将面向全极化 SAR 图像的 GD-Wishart 方法和面向单极化 SAR 图像的 HDEC-TFA 方法得到的散射类别空间分别记为 \mathcal{Y}^1 和 \mathcal{Y}^2 。设 GD-Wishart 法共生成了 y^1 种极化散射类型，HDEC-TFA 共生成 y^2 种子带散射类型，一般情况下 $y^1 \neq y^2$ 。在这里，将从 SAR 图像数据到散射分类结果图的过程看做一个信息传输系统，SAR 图像位于 (x_k, y_k) 的目标信息通过系统传输到散射标签中，对全极化 SAR 图像，目标信息对应极化矩阵 S ，而对于单极化 SAR 图像，目标信息则对应子带散射图 r 。

对极化散射类别 $i \in \mathcal{Y}^1$ 和子带散射类别 $j \in \mathcal{Y}^2$ ，类别 i 中包含 j 越多的信息（或类别 j 中包含 i 越多的信息），则说明极化散射类别 i 和子带散射类别 j 之间

具有越深的联系。因此，本节提出用逐点互信息（Point-wise Mutual Information, PMI）来定量描述两种类别之间的信息传输关系， i 和 j 类之间的 PMI 定义为

$$\text{PMI}_{Y^1 Y^2}(i, j) = \log \frac{P_{Y^1 Y^2}(i, j)}{P_{Y^1}(i) P_{Y^2}(j)} = \log \frac{P_{Y^1|Y^2}(i, j)}{P_{Y^1}(i)}, \quad \dots (5.20)$$

其中 $P_{Y^1 Y^2}(i, j)$ 表示类别 i 和 j 的联合分布， $P_{Y^1}(i)$ 和 $P_{Y^2}(j)$ 分别表示类别 i 和 j 的边缘分布， $P_{Y^1|Y^2}(i, j)$ 表示条件分布。直观的解释如下，如果极化散射类别 i 和子带散射类别 j 是完全互斥的，则 $P_{Y^1|Y^2}(i, j) = 0$ ，而对应的 $\text{PMI}_{Y^1 Y^2}(i, j) = -\infty$ 。如果极化散射类别 i 和子带散射类别 j 之间是相互独立，没有任何联系的，那么 $P_{Y^1 Y^2}(i, j) = P_{Y^1}(i) P_{Y^2}(j)$ ，对应的 $\text{PMI}_{Y^1 Y^2}(i, j) = 0$ 。由于极化分析是固定的物理过程，因此极化散射类别的边缘分布 $P_{Y^1}(i)$ 保持不变，类别 i 和 j 之间的 PMI 随着条件分布 $P_{Y^1|Y^2}(i, j)$ 的增大而增大，因此当条件分布 $P_{Y^1|Y^2}(i, j)$ 达到最大值时，极化散射类别 i 和子带散射类别 j 所表示的物理特性具有最大的正向相关性。

计算所有可能的 $\text{PMI}_{Y^1 Y^2}(i, j)$ 可以得到逐点互信息矩阵，理解上可以类比监督分类结果评估时采用的混淆矩阵（Confusion Matrix）。为了计算 $\text{PMI}_{Y^1 Y^2}(i, j)$ ，需要知道 $P_{Y^1 Y^2}(i, j)$ 、 $P_{Y^1}(i)$ 和 $P_{Y^2}(j)$ 的经验性估计，这里用归一化的计数矩阵 M 来表示，即

$$M(i, j) = \frac{1}{N} \sum_{k=1}^N \delta_{(i,j)}(y_k^1, y_k^2), \quad \dots (5.21)$$

其中

$$\delta_{(i,j)}(y_k^1, y_k^2) = \begin{cases} 1 & y_k^1 = i \text{ and } y_k^2 = j \\ 0 & \text{others} \end{cases} \quad \dots (5.22)$$

由此，联合分布和边缘分布的经验性估计可以分别定义为

$$\hat{P}_{Y^1 Y^2}(i, j) = M(i, j), \quad \dots (5.23)$$

$$\hat{P}_{Y^1}(i) = \sum_j M(i, j), \quad \dots (5.24)$$

$$\hat{P}_{Y^2}(j) = \sum_i M(i, j). \quad \dots (5.25)$$

根据公式 (5.20)，可以得到 PMI 矩阵的经验性估计 $\widehat{\text{PMI}}_{Y^1 Y^2}(i, j)$ 。由于只有 PMI 非负的情况下才表明极化散射类别和子带散射类别中包含的物理特性信息有所

重叠，有利于分析二者之间的关系，因此在计算过程中将 PMI 为负值的情况统一处理为 0。

完成第一层的训练后，对每种子带散射类别进行可视化，并计算 PMI 矩阵 $\text{PMI}^{(1)}$ 进行评估。为了选择更有意义的“平稳类”子带散射类别进行第二轮的精细化聚类分析，这里根据 PMI 矩阵提供的定量化分析结果制定了以下规则。对 DEC 算法中得到的子带散射类别 $j \in \mathcal{Y}^2$ ，设置一个阈值 $\lambda = 1$ ，如果在极化散射类别 \mathcal{Y}^1 中有不止一类的散射特性 i ，使得 $\text{PMI}_{\mathcal{Y}^1\mathcal{Y}^2}(i, j) > \lambda$ ，则子带散射类别 j 被作为“平稳类”选中进入下一轮训练。

结束第二层的训练后，将最后的分类结果可视化并计算 PMI 矩阵 $\text{PMI}^{(2)}$ 。

5.3.5 颜色编码

为了生成最后的子带散射特性分类结果图，使其和极化散射特性结果图一样符合人眼的视觉理解，本节基于 PMI 矩阵的定量化评估结果设计了颜色编码的几条规则。

在极化散射特性结果图的可视化中，用蓝色系、红色系、绿色系和棕色系分别表示奇次散射、二次散射、体散射和混合散射。本节基于定量化的 PMI 矩阵，和 HSI 颜色模型来决定各类子带散射类别的颜色。

- 对“非平稳类”的第一层散射类别 $j^{(1)}$ ，如果存在有且只有一个 i ，使得 $\text{PMI}^{(1)}(i, j^{(1)}) > \lambda$ ，则将 $j^{(1)}$ 类子带散射类别的色调（Hue, H）定为极化散射类别 i 对应的颜色；由此保证了相关性较强的极化散射特性和子带散射特性在视觉上较为相似。

- 对“非平稳类”的第一层散射类别 $j^{(1)}$ ，如果不存在 i 使得 $\text{PMI}^{(1)}(i, j^{(1)}) > \lambda$ ，则说明该子带散射类别与极化散射特性的关联性较差，因此可选取 $\text{PMI}^{(1)}(:, j^{(1)}) > \lambda$ 值最大时对应的极化类别 i 的色调，并将亮度（Intensity, I）调低。

- 对“平稳类”的散射类别 $j^{(1)}$ ，选取 $\text{PMI}^{(1)}(:, j^{(1)}) > \lambda$ 值最大时对应的极化类别 i 的色调作为主色调，第二层的训练结果对应的颜色由第一层类别决定的主色调以及 $\text{PMI}^{(2)}(:, j^{(2)})$ 的值共同决定，可选择在亮度和饱和度（Saturation, S）上进行调整。

5.4 实验与分析

5.4.1 实验数据与实验设计

虽然本章提出的方法针对高分辨率的单极化 SAR 图像，但为了与极化分析的结果进行对比，因此选取了高分三号（GF-3）(Sun 等, 2017) 的全极化 SAR 数据进行实验。

高分三号是一颗工作在 C 波段的多极化合成孔径雷达卫星，一共可提供 12 种类型的观测模式，涵盖单极化、双极化和全极化数据。本节实验选取了 GF-3 在全极化条带模式下对巴黎和旧金山两个城市的成像结果（复数格式）(Sun 等)，名义分辨率 8 米，具体的参数等信息如表 5.1 所示。实验在巴黎和旧金山两景 SAR 图像中分别选取了 2854×3120 和 2048×1750 大小的城市区域进行分析，在地图中的具体位置以及对应的全极化伪彩色图展示在图 5.6 中。

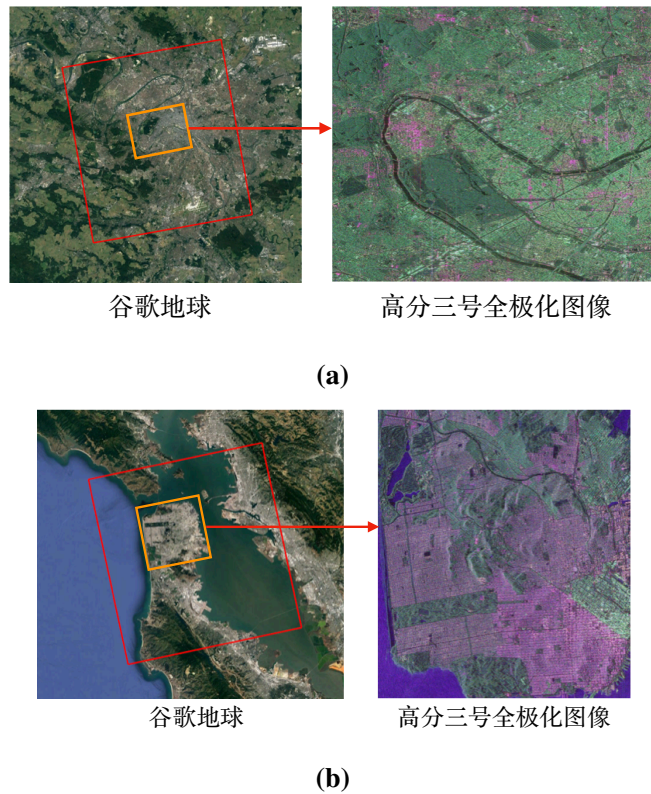


图 5.6 高分三号全极化实验数据。(a) 巴黎，(b) 旧金山。

Figure 5.6 The experimented GF-3 full-polarized SAR scenes. (a) Paris, France, (b) San Francisco, USA.

复数 SAR 图像的时频分析的窗口大小设为 32×32 ，连续子带分解时所用的频域滤波器带宽是原始全孔径图像带宽的一半。在极化分析计算 \mathbf{K} 矩阵，以及

表 5.1 高分三号实验数据信息

Table 5.1 The description of our experimented GF-3 data.

城市	旧金山	巴黎
卫星	高分三号	高分三号
波段	C	C
成像模式	全极化条带	全极化条带
极化	HH/HV/VH/VV	HH/HV/VH/VV
实验区域大小 (像素)	2048×1750	2854×3120
分辨率-距离向 (m)	8	8
分辨率-方位向 (m)	6~9	6~9
像素间隔-距离向 (m)	2.25	2.25
像素间隔-方位向 (m)	5.36	5.3
入射角 (°)	19~22	35~37

HDEC-TFA 方法计算子带散射图时,均采用 3×3 的窗进行平均。

在进行 HDEC 学习时,需要考虑聚类中心个数 k 的选择。[Spigai 等 \(2011\)](#); [Tupin 和 Tison \(2004\)](#) 提出四种典型的子带散射类型,考虑到实际 SAR 场景中散射的复杂程度,在第一层的 DEC 学习时聚类中心 k 应大于 4. 选取 512×512 大小的 SAR 图像切片对不同的 k 取值进行第一层的 DEC 学习,结果如图5.7所示, k 从 5 依次增加到 10 的情况下,随着类别数量增加,与人造目标相关的子带散射类型趋于稳定 (例如第一列和第二列),而多与自然目标相关的子带散射类型开始分裂成更多的小簇 (例如第五列和第七列)。由于人造目标比较特殊的物理性质,例如共振特性、非平稳特性等,其子带散射图反映了更多有关目标物理特性的附加信息;而对多数自然目标来说,子带散射图并不能带来更多有用的散射信息。因此实验最终在第一阶段的 DEC 学习过程中选择了 k 等于 6 或 7,第二阶段训练时 k 等于 3。

5.4.2 可视化与量化分析

本节给出可视化的实验结果,以及量化的 PMI 矩阵分析结果。考虑到巴黎和旧金山两个城市存在差异性,因此本节对两个地区进行了独立的训练和测试,形成不同类别的散射特性。

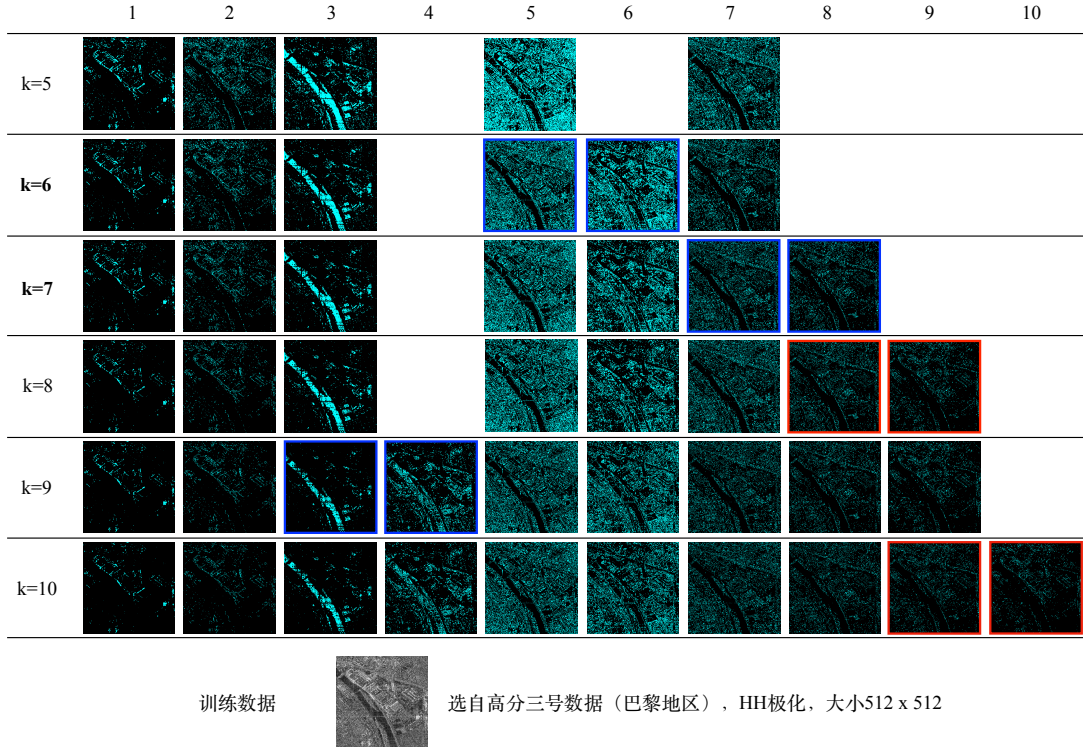
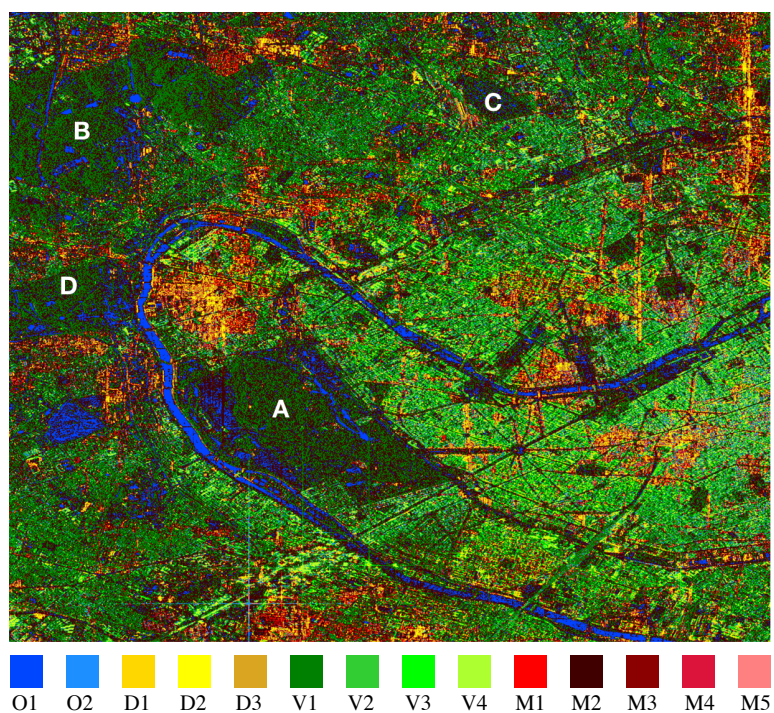


图 5.7 对于 512×512 大小的 SAR 图像切片选取不同 k 值时 DEC 的训练结果

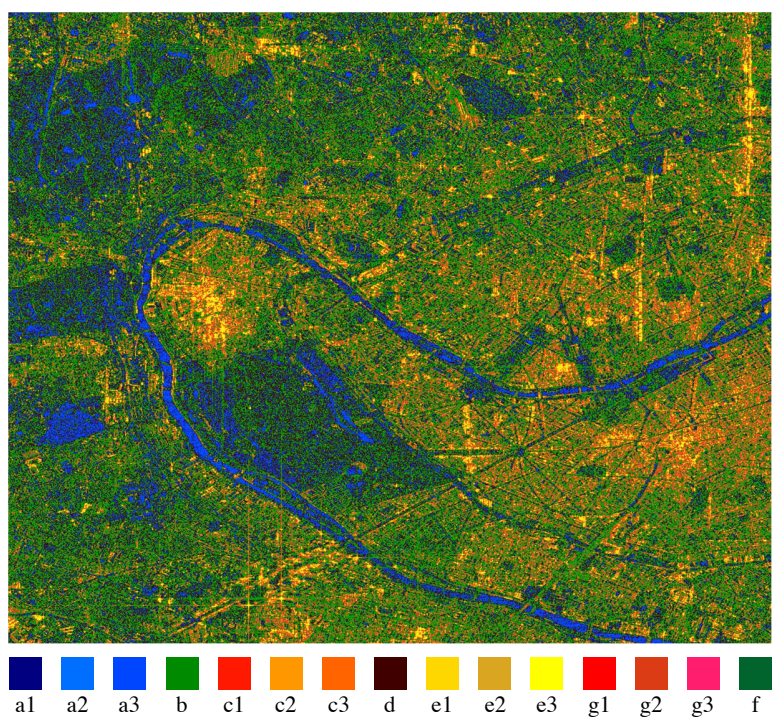
Figure 5.7 The first level DEC training for an area of 512×512 pixels with different cluster number k .

巴黎地区的全极化 SAR 图像的极化散射特性分类可视化结果如图5.8(a) 所示，GD-Wishart 方法将巴黎地区的全极化 SAR 图像共分成了 14 种散射类别，包括两种类型的奇次散射（蓝色系）、三种类型的二次散射（黄/棕色系）、四种类型的体散射（绿色系）以及五种类型的混合散射类（红色系），记为 $\mathcal{Y}^1 = \{O1, O2, D1, D2, D3, V1, V2, V3, V4, M1, M2, M3, M4, M5\}$ ， $y^1 = 14$ 。在图5.8(a) 中可以观察到， $O1$ 代表非常典型的单次散射，一般出现在水体和平整的土地区域；体散射分布在茂密的林区和密集的高楼地区，其中 $V1$ 代表茂密林区的体散射， $V3$ 和 $V4$ 主要出现在不同密度的高楼地区。在混合散射类别中， $M2$ 在自然物区域出现，而 $M1$ ， $M4$ 和 $M5$ 主要分布在城市区域，与二次散射类别 D_i 表达的散射特性较为相似。

对 HH 通道的单极化的 SAR 图像，第一阶段的 DEC 训练选择 k 值为 7，即 $\mathcal{Y}_{level-1}^2 = \{a, b, c, d, e, f, g\}$ ， $y_{level-1}^2 = 7$ ；根据上节设定的规则，类别 a 、 c 、 e 和 g 被挑选出来进入第二阶段的训练，最终得到 15 类的子带散射类型，即



(a)



(b)

图 5.8 巴黎地区。(a) 全极化 SAR 的极化散射特性分类结果图，(b) 单极化 SAR 的子带散射特性分类结果图。

Figure 5.8 The MI matrix in San Francisco area. (a) Level-1 training, (b) Level-2 training.

$\mathcal{Y}_{level-2}^2 = \{a1, a2, a3, b, c1, c2, c3, d, e1, e2, e3, f, g1, g2, g3\}$, $y_{level-2}^2 = 15$. 最终的子带散射特性无监督分类结果如图5.8(b)所示, 与图5.8(a)中所展示的全极化 SAR 的极化散射特性分类可视化结果对比, 在视觉上具有一定的相似性。

旧金山地区的实验结果展示在图5.9中。在全极化 SAR 的 GD-Wishart 分类结果中, 共形成了 12 类极化散射类型, 包括 5 种奇次散射 (蓝色系)、两种二次散射 (黄色系)、两种体散射 (绿色系) 以及三种混合散射类型 (红色系), 即 $\mathcal{Y}^1 = \{O1, O2, O3, O4, O5, D1, D2, V1, V2, M1, M2, M3\}$, $y^1 = 12$. 从图5.9(a)中可以观察到, 奇次散射类别中的 $O4$ 和 $O5$ 出现在水域 (如河流、湖泊和海洋) 和一些平整的地面区域, 而 $O1$, $O2$ 和 $O3$ 都主要发生在城市区域, 对应一些人造目标。对于体散射类别, 很容易看出 $V1$ 表示高密度高楼区域的多次散射, 而 $V2$ 代表植被树林区域的体散射。二次散射 Di 和混合散射 Mi 基本上都出现在居民区, 对应房屋等人造目标。

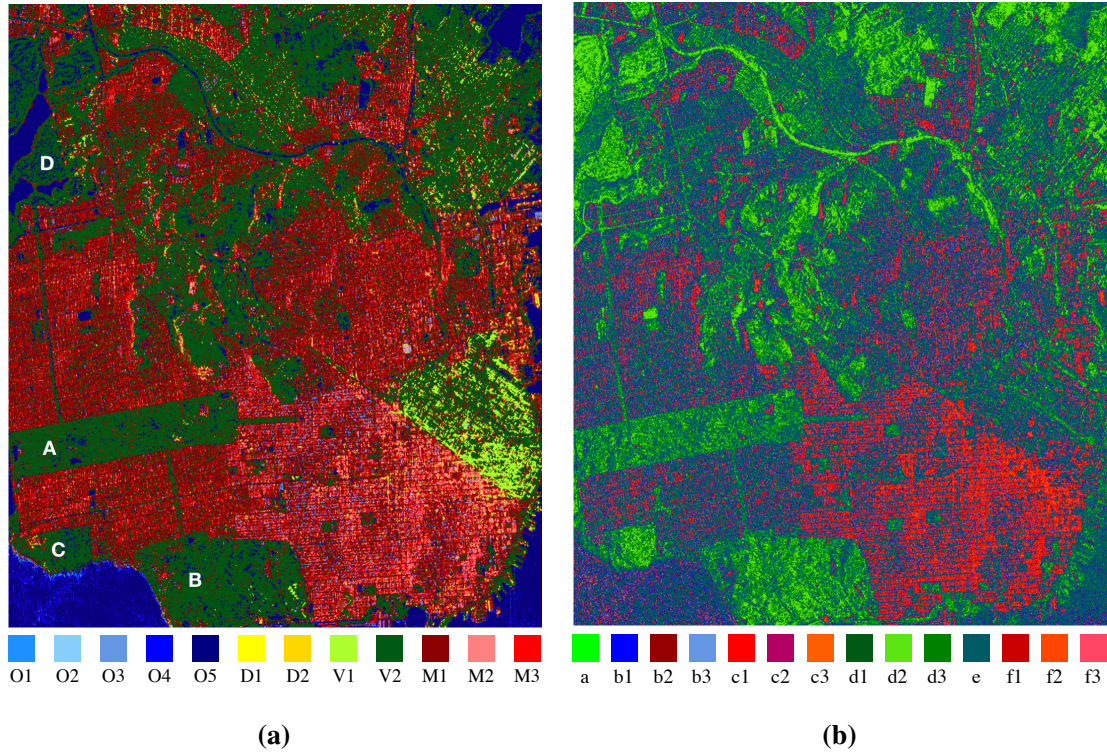


图 5.9 旧金山地区。(a) 全极化 SAR 的极化散射特性分类结果图, (b) 单极化 SAR 的子带散射特性分类结果图。

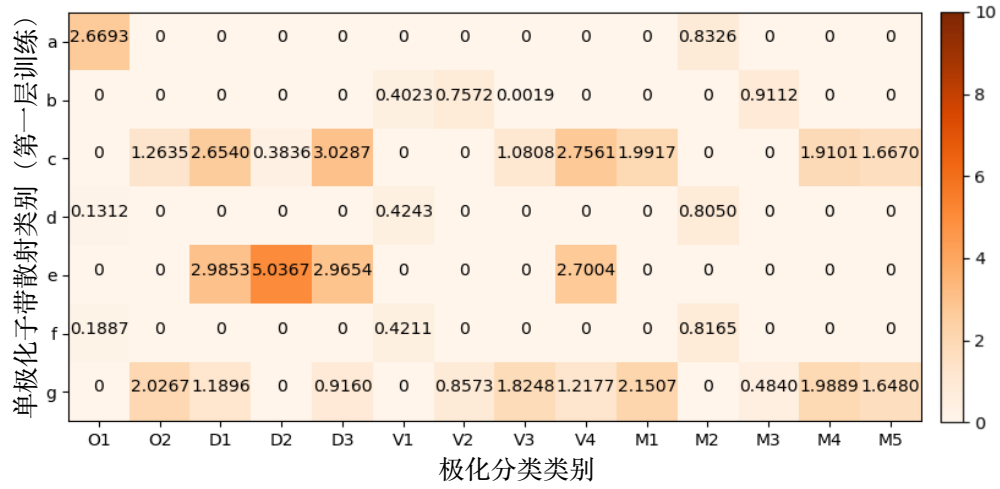
Figure 5.9 The visualization result in San Francisco area. (a) The polarimetric scattering mechanisms for Quad-Pol SAR image, (b) The sub-band scattering behaviours for single-polarized SAR images.

旧金山城区的分布相比于巴黎呈现出更简单的城市结构,在极化散射特性分析结果中的散射类别相比巴黎地区更少,因此在对单极化 SAR 数据应用 HDEC-TFA 方法时,第一阶段的训练选取了 k 值为 6,即 $\mathcal{Y}_{level-1}^2 = \{a, b, c, d, e, f\}$, $y_{level-1}^2 = 6$ 。然后类别 b, c, d 和 f 输入到第二阶段的 DEC 训练,生成最终 14 类子带散射类型,即 $\mathcal{Y}_{level-2}^2 = \{a, b1, b2, b3, c1, c2, c3, d1, d2, d3, e, f1, f2, f3, g\}$, $y_{level-2}^2 = 14$ 。最终的可视化结果如图5.9(b)所示。

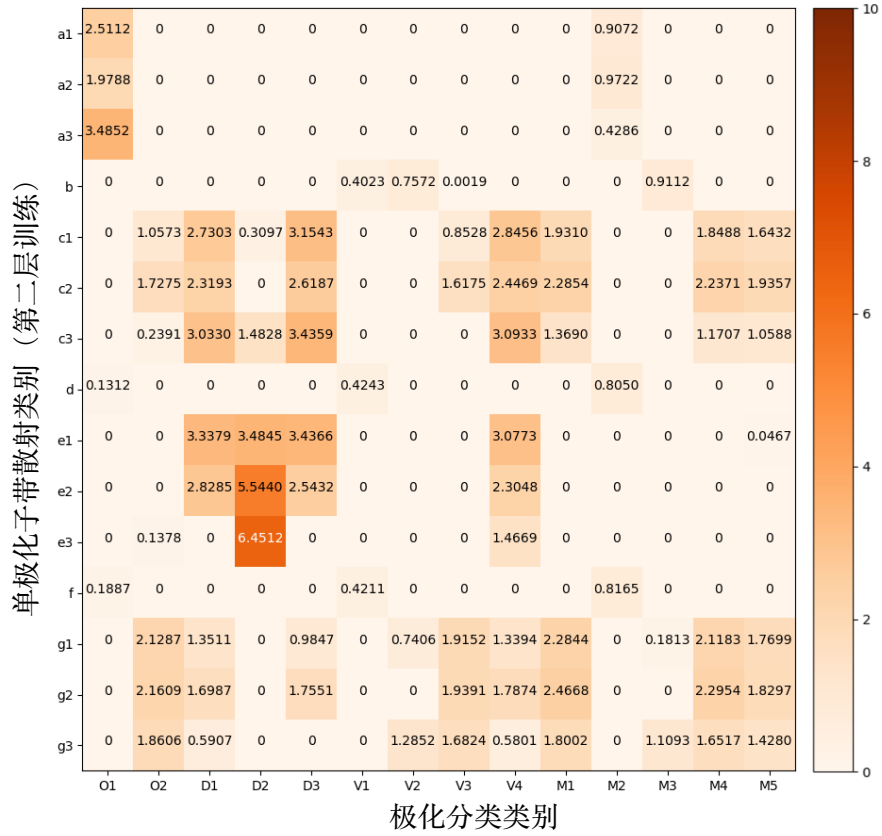
下面对两个地区的结果进行定量化的分析与讨论。

首先以巴黎地区的结果为例,全极化 SAR 的 GD-Wishart 方法共生成 $y^1 = 14$ 类极化散射类别,在第一阶段的训练中选取 $y_{level-1}^2 = 7$,则计算出第一阶段 DEC 训练结果的 PMI 矩阵,大小为 14×7 ,如图5.10(a)所示。最大值 $\text{PMI}^{(1)}(D2, e) = 5.0367$ 表明子带散射类别 e 与极化分析的二次散射 $D2$ 类别关联性最强,接下来的 $\text{PMI}^{(1)}(D1, e) = 2.9853$, $\text{PMI}^{(1)}(D3, e) = 2.9654$ 和 $\text{PMI}^{(1)}(V4, e) = 2.7004$ 表明子带散射类别 e 与极化分析的其他两个二次散射类别 $D1$ 和 $D3$,以及体散射类别 $V4$ 也具有较强相关性。第一行的 PMI 值中,只有 $\text{PMI}^{(1)}(O1, a) = 2.6693$ 是大于 $\lambda = 1$ 的,其他情况的 PMI 值都较小或为 0,由此说明子带散射类别 a 与单次散射 $O1$ 表现的物理特性具有最大的联系。子带散射类别 c 和 g 对应的多个 $\text{PMI}^{(1)}(:, c)$ 和 $\text{PMI}^{(1)}(:, g)$ 都具有大于 λ 的互信息。而 b, d 和 f 类的 PMI 值都小于 λ ,被判定为“非平稳类”子带散射类别。因此最终将第一阶段的 a, c, e 和 g 四类散射类别送到第二阶段的训练中,每类散射特性又继续分裂成三个小类。

最终的分类结果一共包括 15 种子带散射类别,因此第二层训练完毕后得到了大小为 14×15 的 PMI 矩阵,如图5.10(b)所示。这里观察到一个现象,即第一阶段训练得到的某些子带散射类别在经过第二阶段的训练后,其分裂的子类具有更独特的散射特性,并且与 GD-Wishart 方法得到的极化散射特性的联系更大。例如 $\text{PMI}^{(2)}(D2, e3) = 6.4512$ 与 $\text{PMI}^{(2)}(V4, e3) = 1.4669$ 和 $\text{PMI}^{(2)}(O1, e3) = 0.1887$ 相比,体现出子带散射类别 $e3$ 与二次散射类别 $D2$ 的关联性更强。因此,采用多层级的训练方式使得子带散射分类结果更具代表性,特别是对于散射较复杂的人造目标,多层级的 HDEC-TFA 方法揭示了更多有意义的子带散射特性,在定量分析中表现出提升主要的 PMI 值,抑制次要的 PMI 值,与某种明确的极化散射特性的关联性更强。例如 $\text{PMI}^{(2)}(O1, a3) = 3.4582$ 相比 $\text{PMI}^{(1)}(O1, a) = 2.6693$ 提升了 0.8159,而对应的 $\text{PMI}^{(2)}(M2, a3) = 0.4286$ 相比于 $\text{PMI}^{(1)}(M2, a) = 0.8326$



(a)



(b)

图 5.10 互信息矩阵 (巴黎地区)。(a) 第一层训练, (b) 第二层训练。

Figure 5.10 The MI matrix in Paris area. (a) Level-1 training, (b) Level-2 training.

减小了 0.404.

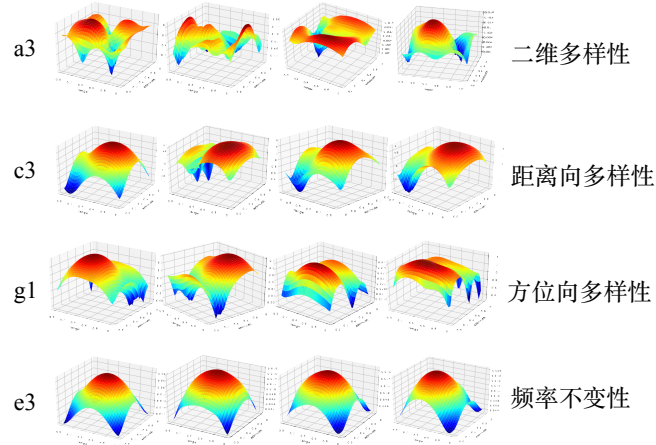
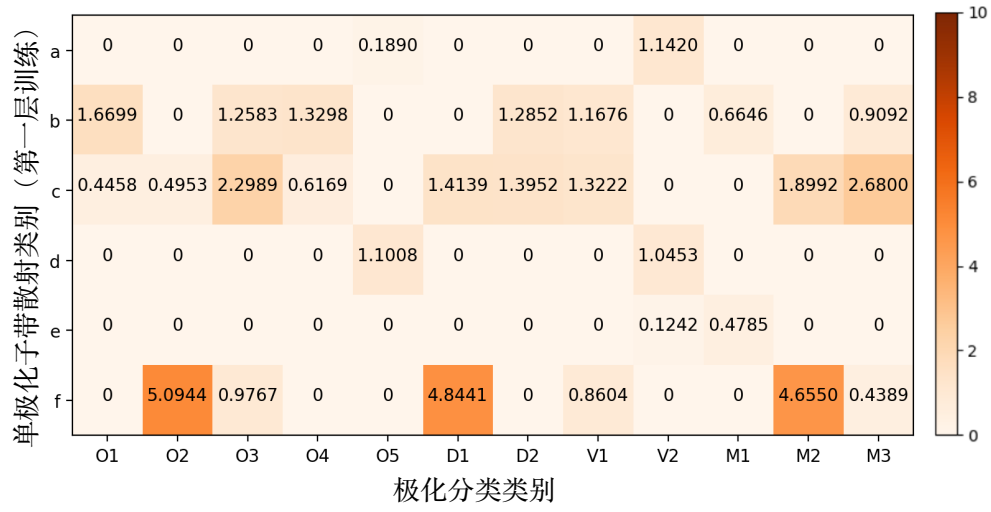


图 5.11 子带散射类别 $a3$, $c3$, $e3$ 和 $g1$ 的子带散射图示例

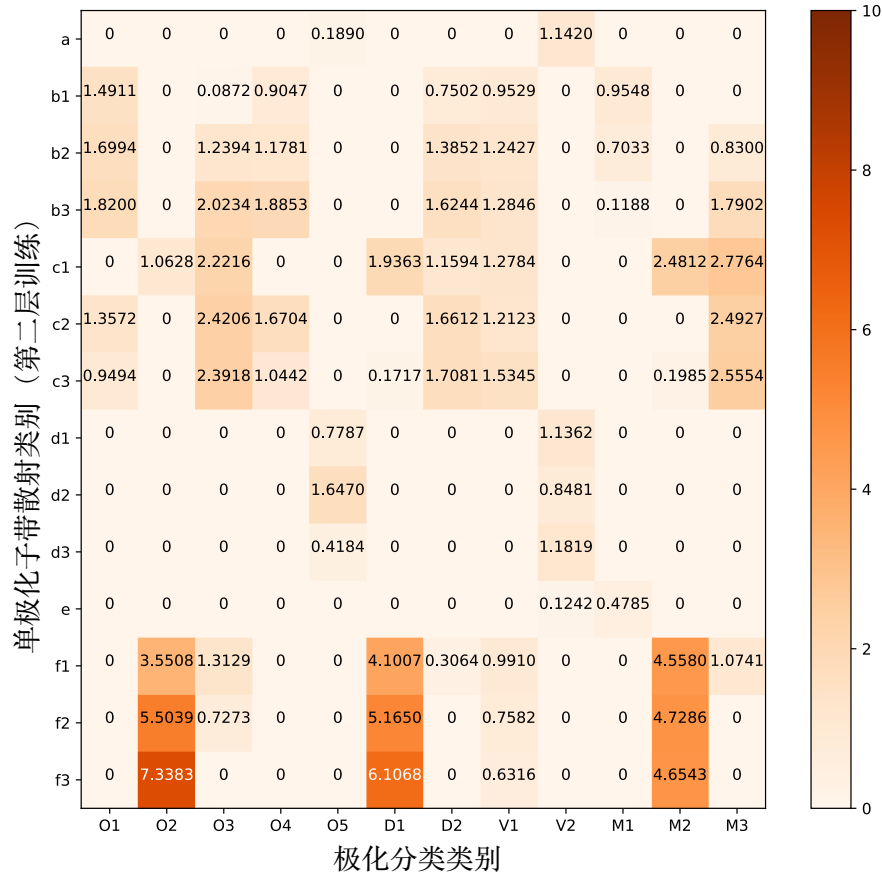
Figure 5.11 Some examples for sub-band scattering patterns of $a3$, $c3$, $e3$ and $g1$ classes.

在最终的子带散射分类结果中, 挑选出 a , c , e 和 g 四种子带散射类型中 PMI 值最高的子类 $a3$, $c3$, $e3$ 和 $g1$ 做进一步的分析和讨论。 $a3$, $c3$, $e3$ 和 $g1$ 四个子类分别与极化散射类别 $O1$, $D3$, $D2$ 和 $M1$ 的关系最紧密, 因此图5.11中给出了四个子带散射类型的一些子带散射图的示例, 基本能与第5.3.2节描述的四种基础散射模式一一对应。通过分析 PMI 矩阵, 可以发现子带散射类别 $a3$ 与极化分析中的单次散射 $O1$ 联系最紧密, 其 PMI 值为 $\text{PMI}^{(2)}(O1, a3) = 3.4582$, 图5.11中展示的子带散射图例子均表现出二维多变性的特征, 散射强度在距离向和方位向上出现无规律的波动, 多在二维杂波信号中出现, 一般对应一些自然地物。从图5.8(b) 中可以看到子带散射类别 $a3$ 出现在河流和裸露地面区域。从极化分析的角度, 这些地方一般体现出表面散射的特性, 因此在极化散射特性分类结果中, 单次散射类别 $O1$ 与子带散射类别 $a3$ 联系最紧密。对于具有频率不变性的各向同性目标, 在距离向和方位向上的散射强度较为稳定, 如 $e3$ 中的几个子带散射图所示, 粗糙地面上的二面角或三面角结构会反映出这样的性质, 在极化分析中体现出二次散射的特性。因此图5.10(b) 所示的 PMI 矩阵中, 互信息值最高的 $\text{PMI}^{(2)}(D2, e3) = 6.4512$ 表明子带散射类别 $e3$ 与二次散射 $D2$ 具有最强的相关性。

旧金山区域实验结果的 PMI 矩阵如图5.12所示, 根据 $\text{PMI}^{(2)}(O2, f) = 5.0944$, $\text{PMI}^{(2)}(D1, f) = 4.8441$ 以及 $\text{PMI}^{(2)}(M2, f) = 4.6550$ 可以得出, 子带散射类别 f



(a)



(b)

图 5.12 互信息矩阵 (旧金山地区)。(a) 第一层训练, (b) 第二层训练。

Figure 5.12 The MI matrix in San Francisco area. (a) Level-1 training, (b) Level-2 training.

与多种极化散射类别如 $O2$, $D1$ 和 $M2$ 相关联。图5.13展示了极化分析得到的混合散射类别 $M2$ 和单极化 SAR 图像的子带散射类别 f 的可视化结果图, 可以观察到两种类别的一些共性。在极化分析结果中, 奇次散射 $O2$ 和二次散射 $D1$ 发生在一些特殊的人造目标上, 例如房顶和船坞。图5.14给出的两个典型目标的分析结果, 以及 PMI 矩阵中 $(O2, f)$ 和 $(D1, f)$ 展示出的较高 PMI 值, 均印证了子带散射类别 f 与 $O2$ 和 $D1$ 均有着密切的联系。

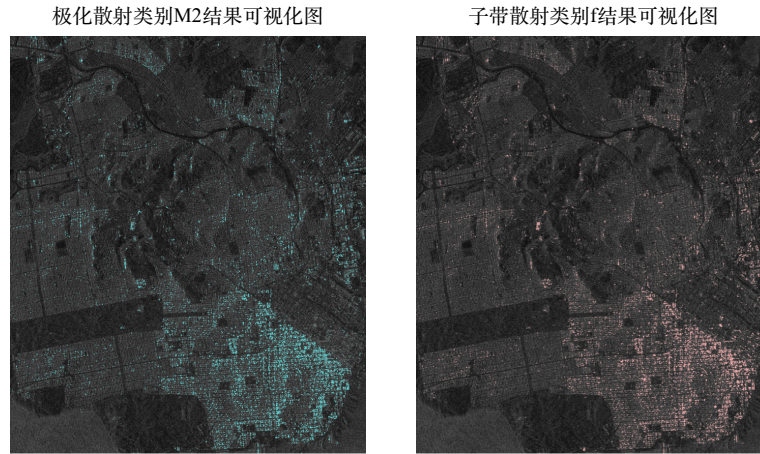


图 5.13 旧金山地区极化散射特性 $M2$ 和子带散射类别 f 的结果可视化图

Figure 5.13 The class f in HDEC-TFA result and class $M2$ in GD-Wishart result, San Francisco.

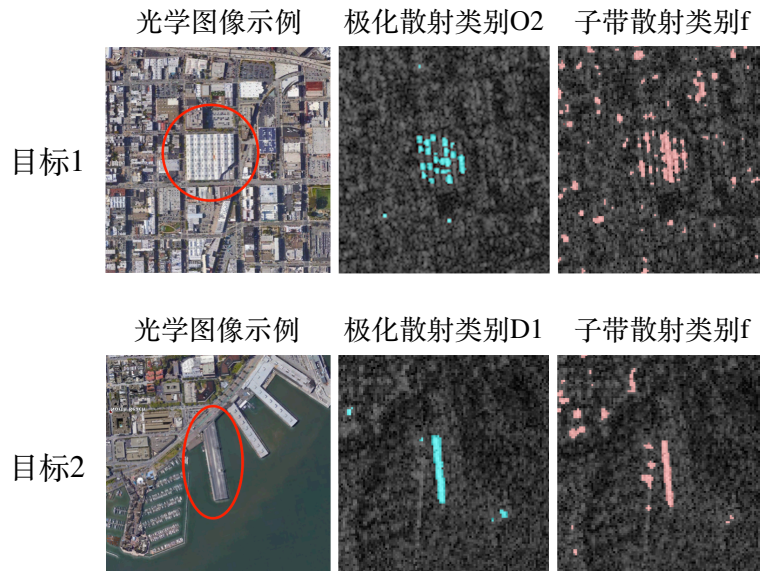


图 5.14 旧金山地区两个目标的子带散射类别 f 和极化散射特性 $O2$ 及 $D1$ 的结果可视化图

Figure 5.14 For two particular targets in San Francisco, HDEC-TFA result of class f and GD-Wishart result of class $O2$ and $D1$ are shown.

5.4.3 具体目标和场景分析

本节将对一些特殊的人造目标和具体的自然场景进行分析和讨论。

首先，图5.15展示了旧金山地区的目标分析结果，A、B和C标记了三种建筑的屋顶，都具有相似的物理结构，即屋顶与地面呈锐角，且朝向相同。实验采用的高分三号数据入射角在 $19^{\circ}\sim 22^{\circ}$ 之间，对于这种结构一般呈现出单次散射特性。从图5.15中能得到一部分印证，即目标A和部分目标B都表现出奇次散射类别 $O2$ ，然而目标C的散射类别在极化分析结果中比较模糊，可能会受到目标C附近植被等因素的影响，无法体现出单次散射的特性。另一方面，对单极化 SAR 图像进行的 HDEC-TFA 分析结果显示，目标A、B和C都能得到很清晰的子带散射特性提取结果。尽管单极化 SAR 图像缺乏极化信息，但实验表明，与面向全极化 SAR 图像的极化分析方法相比，本章所提出的 HDEC-TFA 方法在对一些结构特殊的人造目标进行散射特性分析时能够体现出一些优势，有利于在高分辨率的单极化 SAR 图像中对密集复杂的城市区域提供丰富的有意义的物理散射知识。

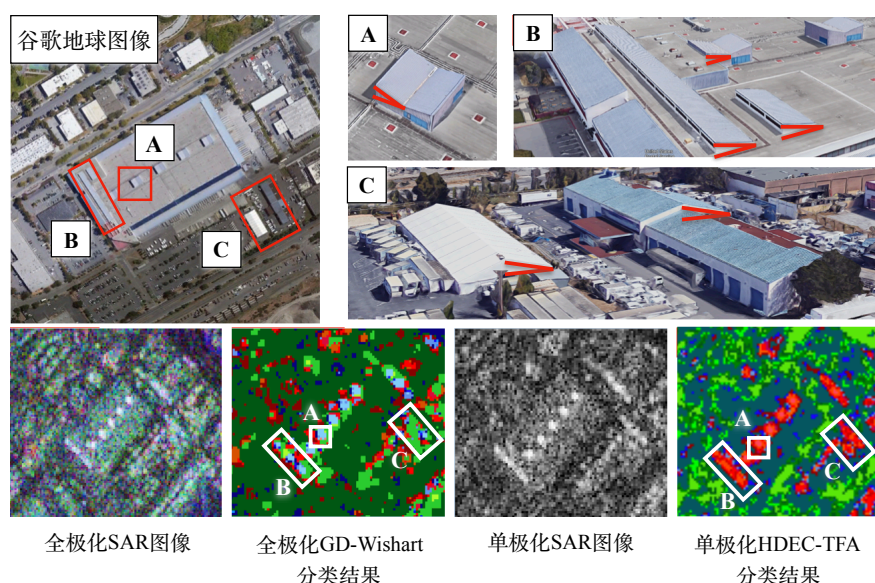


图 5.15 旧金山地区人造目标散射特性结果分析

Figure 5.15 The scattering mechanism analysis for man-made targets in San Francisco.

图5.16和图5.17给出了巴黎地区的几个人造目标，例如埃菲尔铁塔、火车站、街区以及高楼的散射特性分析结果。与面向全极化 SAR 数据的极化分析方法 GD-Wishart 相比，只针对单极化 SAR 图像的非监督 HDEC-TFA 方法在特别的人

造目标的物理散射特性提取上能达到类似的效果，依托不同的理论，HDEC-TFA方法从另一个角度揭示了目标的物理散射特性。

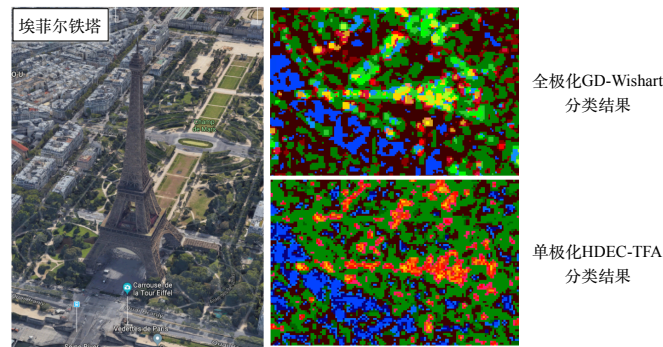


图 5.16 巴黎埃菲尔铁塔散射特性结果分析

Figure 5.16 The scattering mechanism analysis for Eiffel Tower in Paris.

从图5.10和图5.12的 PMI 矩阵中可以发现，有几类子带散射类别与极化散射类别的联系非常弱，通过分析发现这些类别大多对应了一些自然地物，例如草地、丛林等。前面的讨论主要都集中在与极化散射类别联系较紧密的子带散射特性，下面将主要对联系较弱的几类子带散射类别展开进一步的分析，讨论极化分析结果和子带散射类别在 PMI 值较低的情况下是否还存在潜在的联系。

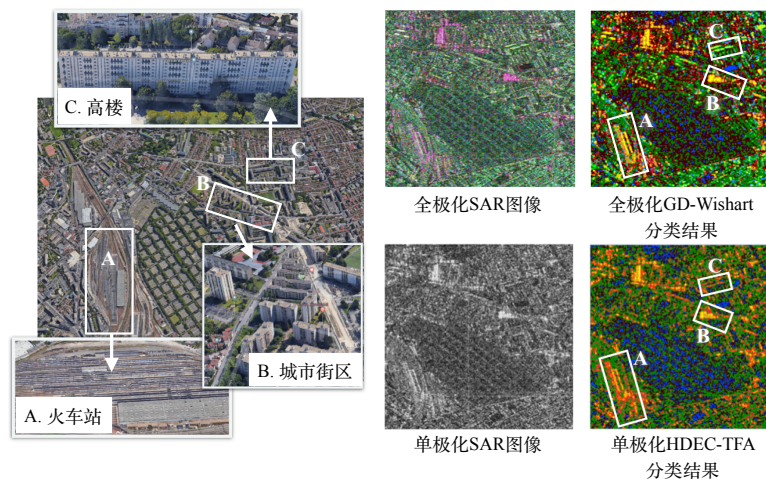


图 5.17 巴黎地区人造目标散射特性结果分析

Figure 5.17 The scattering mechanism analysis for man-made targets in Paris.

图5.8(a) 和图5.9(a) 中分别标记了四个自然场景 A、B、C 和 D，现在在这四个区域分别裁剪大小为 100×100 的小块，对 GD-Wishart 方法和 HDEC-TFA 的结果进行百分比分析，画出每种散射类型占比的饼图，如图5.18和图5.19所示。

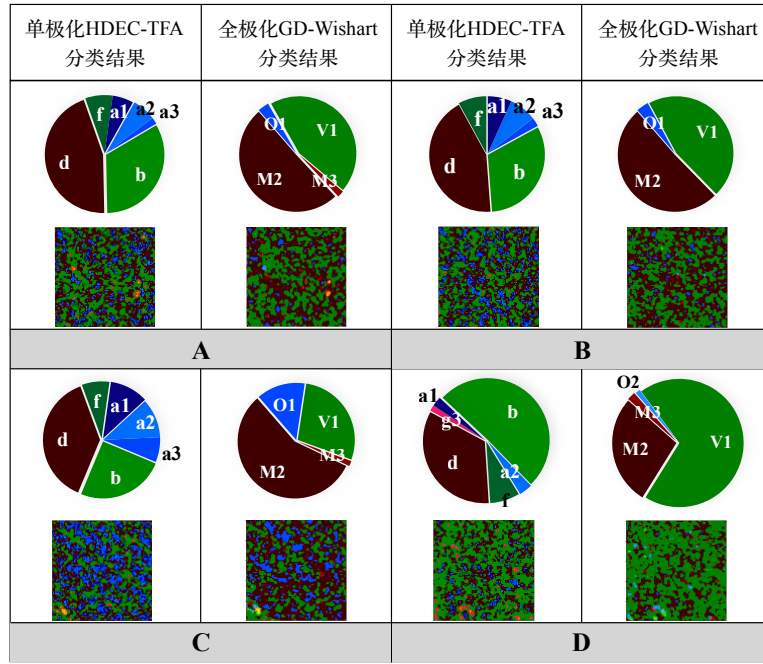


图 5.18 巴黎地区自然地物散射特性百分比分析

Figure 5.18 The percentage analysis of GD-Wishart and HDEC-TFA results for natural area in Paris.

尽管图5.10(a) 第二行的 $\text{PMI}^{(2)}(:, b)$, 第四行的 $\text{PMI}^{(2)}(:, d)$, 以及第六行的 $\text{PMI}^{(2)}(:, f)$ 都很小, 即子带散射特性 b 、 d 、 f 与所有极化散射类别的联系都不大, 从图5.18中仍可以看出, 子带散射类别 b 和 d 在 100×100 的区域中所占百分比与体散射 $V1$ 和混合散射 $M2$ 相似。在旧金山的例子中, 第一层训练后的子带散射类别 d 与奇次散射 $O5$ 和体散射 $V3$ 的互信息相差无几, $\text{PMI}^{(1)}(O5, d) = 1.1008$, $\text{PMI}^{(1)}(V2, d) = 1.0453$, 观察可视化结果也能发现子带散射类别 d 不仅出现在植被区域, 也出现在道路和河流处。为了更好区分不同区域的散射特性, 在第二层的训练中将 d 进行了更细致的分类, 得到 $d1$ 、 $d2$ 和 $d3$ 类。从最终的 PMI 矩阵中可以看到, 子带散射类别 $d2$ 与奇次散射 $O5$ 的 PMI 值增大到 1.6470, 与其他两类 $d1$ 和 $d3$ 相比, 更多地与 $O5$ 类关联, 在可视化结果中主要对应道路和河流区域。而 $d1$ 和 $d3$ 与体散射 $V2$ 的 PMI 值占主要地位, 在可视化结果中主要对应植被区域。这个结论从图5.19中也能得到响应的证明。

5.4.4 数据驱动和理论驱动方法讨论

本章针对全极化 SAR 图像采用的极化分析方法基于由来已久的雷达极化理论, 以及极化 SAR 图像的统计特性, 具有较为坚实的理论基础, 可以广泛应用

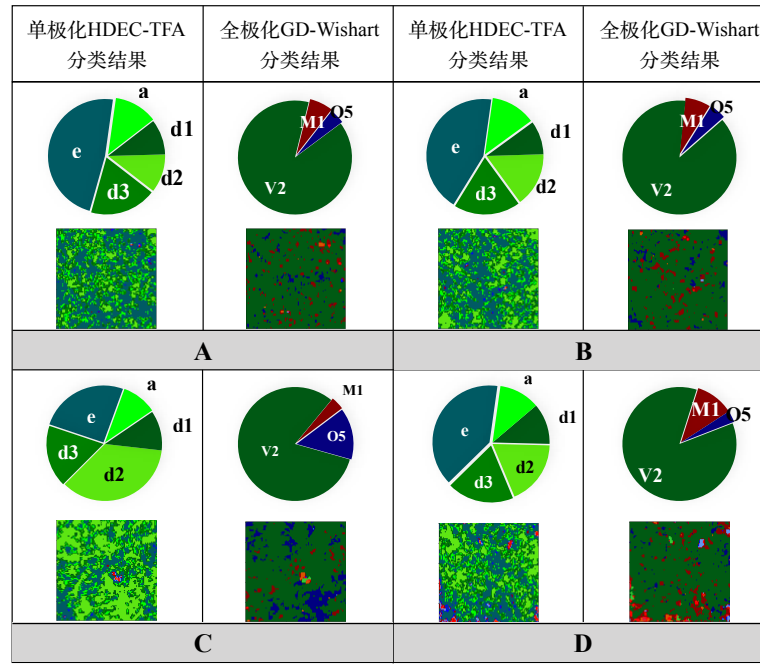


图 5.19 旧金山地区自然地物散射特性百分比分析

Figure 5.19 The percentage analysis of GD-Wishart and HDEC-TFA results for natural area in San Francisco.

到不同的极化 SAR 图像中。当缺乏极化信息时，基于极化理论的分析方法在提取极化特征获取极化散射特性上受到极大限制。为了从单极化 SAR 图像中重建极化特征，Song 等 (2018) 提出一种数据驱动的深度学习学习方法，将单极化 SAR 图像作为输入，全极化 SAR 图像作为真值，设计深度神经网络进行学习，从单通道的 SAR 图像出发重构了彩色极化 SAR 图像并且恢复了全极化 SAR 图像的极化信息。近期，Zhao 等 (2019) 提出用复数卷积神经网络从单极化 SAR 图像中学习极化特征，以全极化 SAR 图像利用 H/α 极化分解得到的极化散射特征作为真值。

本章提出的子带散射特性提取方法与上述两种数据驱动的方法不同，一方面，本章提出的 HDEC-TFA 算法是一种无监督的方法，无需从全极化 SAR 图像中获取反映极化特性的真值。实验分析了 HDEC-TFA 算法与极化分析结果的关联性，只是用于说明本方法学习到的特性具有一定的物理意义。另一方面，HDEC-TFA 算法一部分是理论驱动的，其中的 TFA 部分依托 SAR 图像的时频分析理论提出二维连续子带分解，对目标在距离向和方位向上的散射变化进行描述；一部分又是数据驱动的，其中的 HDEC 部分则是从大量的子带散射图中自

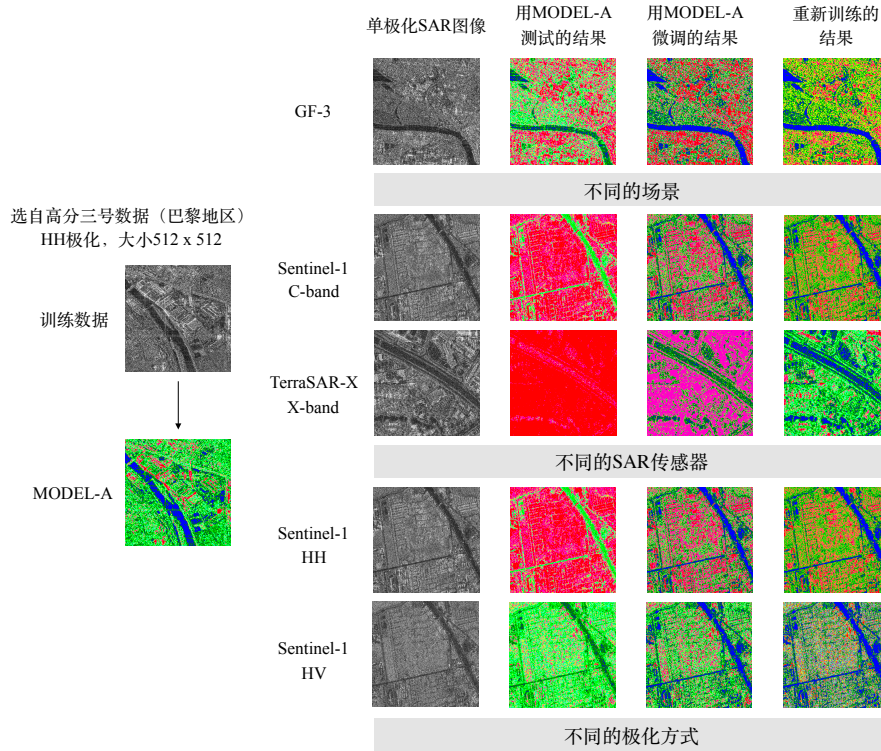


图 5.20 将 MODEL-A 应用到不同 SAR 场景中的实验结果对比

Figure 5.20 The MODEL-A is tested or fine-tuned on SAR images with different scenes, sensors, and polarizations.

动学习特征并进行类别划分。由于理论驱动的 TFA 部分从物理角度描述了目标特性，是复数 SAR 图像中固有的性质，因此 HDEC-TFA 方法可以用在所有单极化复数 SAR 图像中提取子带散射特性。同时数据驱动的特点又使得 HDEC-TFA 方法训练出的深度模型对数据产生依赖性，不一定能适用于其他的 SAR 图像。接下来本节将探讨 HDEC-TFA 训练的模型在不同 SAR 图像数据中的泛化性能。

本节采用图5.7中的数据进行训练，即高分三号巴黎地区的数据中选取的一张 512×512 大小的切片，将其训练的 DEC 模型记为 MODEL-A，如图5.20所示。表5.2展示了实验选取的不同 SAR 图像数据的参数。在本实验中，直接采用 MODEL-A 进行测试的结果和利用 MODEL-A 进行微调的结果与从头训练的结果进行了对比。当直接用 MODEL-A 测试相同条件下不同场景的 SAR 图像时，图5.20中的第一行展现出了一个较能接受的结果，而在进行微调之后，一些目标转移到了更合适的散射类别，例如河流区域，而与人造目标相关的类别（例如红色系的结果）表现得比较稳定。图5.21展示了用 MODEL-A 测试大小为 6144×7680 的巴黎地区高分三号全景 SAR 图像的结果，从可视化角度来看充分

表 5.2 训练源数据与应用的目标 SAR 数据各项参数对比

Table 5.2 The detailed configuration of the trained source data and different target data.

卫星	高分三号 训练源数据	高分三号 不同场景	Sentinel-1 不同极化方式	TerraSAR-X 不同波段
波段	C-band	C-band	C-band	X-band
成像模式	Stripmap	Stripmap	Stripmap	Stripmap
极化方式	HH	HH	HH and HV	HH
分辨率-rg (m)	8	8	2.5	1.18
分辨率-az (m)	6~9	6~9	3.6	3.3
像素间隔-rg (m)	2.25	2.25	2.2	0.9
像素间隔-az (m)	5.36	5.36	3.5	2
入射角 (°)	19~22	19~22	31.2	26.4

说明了模型的潜能。由于训练数据的体量直接与选取 SAR 图像的大小有关，更大的 SAR 图像可以包含更丰富的子带散射类别，进而使得训练的模型泛化能力更强，因此有理由推测采用更多数据训练之后模型在不同场景中的应用更具有灵活性。

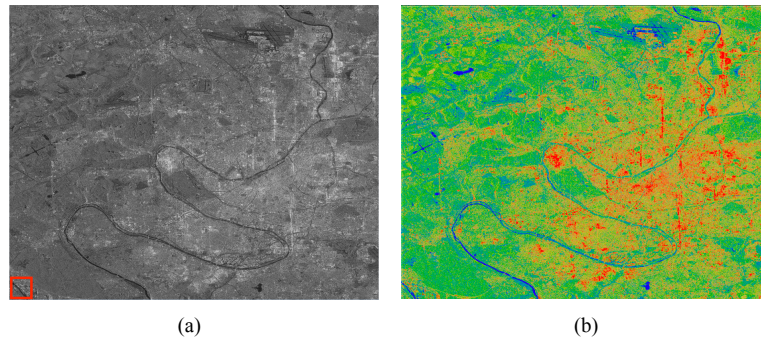


图 5.21 MODEL-A 测试巴黎地区大小为 6144×7680 的全景 SAR 图像的结果。(a)HH 通道 SAR 图像，左下角的红框标记为训练数据，(b)MODEL-A 的测试结果。

Figure 5.21 The result of testing MODEL-A on a SAR scene in Paris with the size of 6144×7680 .

- (a) The HH channel SAR image where the bottom left red rectangle is the training patch.
(b) The result of testing with the trained MODEL-A.

图5.20的第二行和第三行展示了将 MODEL-A 应用到不同 SAR 卫星数据中的实验结果。直接将 MODEL-A 应用于 Sentinel-1 和 TerraSAR-X 卫星数据会出现一些问题，由于基于时频理论提取的子带散射图反映了目标在不同方位角，以

及不同发射频率带宽的散射变化, C 波段的高分三号数据训练的模型即使在微调的情况下也较难应用到 X 波段的图像中。例如 X 波段和 C 波段的 SAR 卫星在电磁波穿透能力上的差异性会导致自然植被地区的散射出现变化, 进而影响到子带散射图的形状, 使网络在特征提取和分类阶段都出现偏差。Sentinel-1 卫星与高分三号一样在 C 波段工作, 相比之下差异较小, 因此在微调之后可以观察到明显的改善。为了进一步验证前面的猜想, 实验还选取了第 5.4.2 小节中大小为 2854×3120 的巴黎地区 SAR 图像训练的模型 (记为 MODEL-P) 对 Sentinel-1 的数据进行测试, 结果如图 5.22 所示, 可以观察到 MODEL-P 在 Sentinel-1 数据上的结果相比于 MODEL-A 的效果更好。除此之外实验还验证了 HH 通道图像训练的模型在 Sentinel-1 的 HV 数据中测试的结果, 如图 5.20 的最后两行所示, 可以观察到在同极化数据的训练模型在同种波段的交叉极化 SAR 图像中也能取得较好的结果。

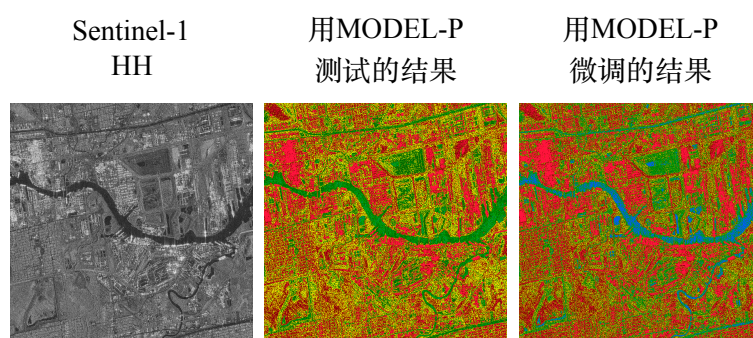


图 5.22 MODEL-P 在 Sentinel-1 SAR 图像上测试和微调的结果

Figure 5.22 Test and fine-tune MODEL-P on Sentinel-1 HH SAR image.

5.5 本章小结

为了对单极化 SAR 图像进行全局的物理散射特性分析, 本章提出了一种无监督的基于时频分析理论的多层级深度嵌入聚类方法 (HDEC-TFA), 基于时频分析理论对每个位置的目标提取子带散射图, 然后通过深度神经网络自动学习子带散射图的潜在特征并自动分类, 得到最终的子带散射特性分类结果图。所提的 HDEC-TFA 方法以数据驱动为主, 目前还缺少足够的理论依据透彻地解释通过无监督学习得到的散射类别, 为了说明学习到的散射特性是有意义的, 本章提出用全极化数据的极化分析结果进行对比论证。结果表明 HDEC-TFA 方法提

取的子带散射特性与极化散射特性有一定的联系，特别是人造目标类别联系较为紧密，并且在某些特殊目标的特性检测上优势明显。这种方法为理解单极化 SAR 图像的物理散射特性提供了新的视角，在缺乏极化信息的情况下也能得到反映散射特性的可视化结果，可以辅助后续 SAR 图像在地物语义层级上的解译。

本章的主要工作已发表于 *IEEE TGRS*，部分工作整理成文被 *EUSAR 2020* 国际会议接收：

- **HUANG Z**, DATCU M, PAN Z, QIU X, LEI B. HDEC-TFA: An Unsupervised Learning Approach for Discovering Physical Scattering Properties of Single-polarized SAR Image [J], **IEEE Transactions on Geoscience and Remote Sensing**, 2020, Early Access
- **HUANG Z**, DATCU M. Urban area analysis in single-polarized SAR images based on unsupervised deep learning [C], **European Conference on Synthetic Aperture Radar (EUSAR)**, 2020

第6章 散射特性约束下的 SAR 图像分类框架

6.1 引言

SAR 图像解译是人类从传感器接收到的回波信号，到地物具体语义类别的一个认知过程。如图6.1所描述的，雷达发射接收电磁波并进行处理形成复数的二维 SAR 图像信号，是一个复杂的物理过程，可以看做传感器对地球表面及目标在信号层级的认知。第三章介绍了一个基于 SAR 幅度图像的深度预训练模型，能提取出抽象的 SAR 图像空间特征，这是机器在视觉层级上对 SAR 图像的理解。视觉层级的理解对于人类来说是直观且友好的，而物理层级的理解隐匿在人类无法“看见”的信号身后，需要辅助的物理模型和理论去支撑和解释。上一章介绍了如何利用深度学习方法从高分辨率的单极化 SAR 图像中自动学习出体现物理散射特性的子带散射类型，与针对极化 SAR 图像的极化分析方法类似，二者都能从复数 SAR 图像中获取散射类型分类结果图，从电磁波和物理散射特性的角度去解释 SAR 图像。而人类最终的认知目标，是从地物的语义层级去理解 SAR 图像，例如划分河流、民居、铁路等语义类别。

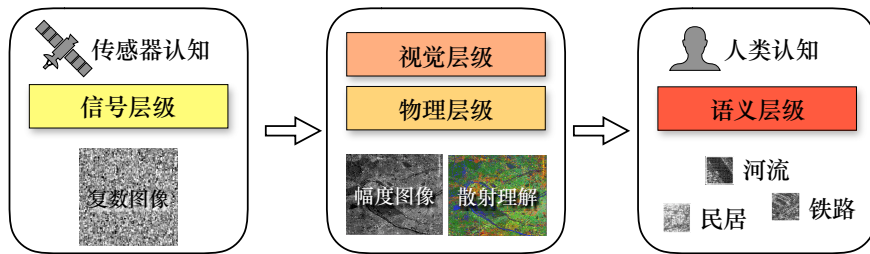


图 6.1 SAR 图像的认知过程

Figure 6.1 The cognitive process of SAR images.

由于目标极化矩阵中包含的丰富的极化信息可以很容易将其与目标的后向散射强度相结合，很多学者面向极化 SAR 图像开展了同时在视觉和物理层级进行 SAR 图像分类的研究。在深度学习驱动的研究中，常采用的策略包括，直接从极化散射矩阵出发设计极化散射编码的神经网络学习 (Liu 等, 2019)，或利用极化散射矩阵中的极化特征代替像素值作为深度神经网络的输入进行学习，例如对每个目标像素点选取极化相干矩阵中元素的实部虚部进行运算和组合，形成不同的多维输入向量 (Zhou 等, 2016b; Chen 和 Tao, 2018; Zhang 等, 2016)；或

者设计基于 Wishart 距离的神经网络来结合深度学习和 PolSAR 分类 (Jiao 和 Liu, 2016) 等。

但在基于深度学习方法的单极化 SAR 图像分类研究中, 很难像极化 SAR 数据一样利用极化散射矩阵或极化信息的特殊分布同时从物理层面和视觉层面进行研究。另一方面, 大规模快速的卫星图像获取带来海量的未知 SAR 数据, 对一个新的 SAR 场景进行理解时已知的信息有限, 如何在标记极少的情况下预测大范围 SAR 场景的语义类别是一个亟待解决的问题。因此本章的研究内容面向两个主要问题展开: 1. 无论是否包含极化信息, 如何结合物理和视觉层面对 SAR 图像进行语义上的区分; 2. 针对极少的标记样本, 如何让神经网络尽可能地学习到更多的知识去理解未知的场景。

第五章为理解单极化的 SAR 图像的物理散射特性提供了一个新视角, 本章基于此提出一个同时面向视觉和物理层级的 SAR 图像分类框架, 针对复数 SAR 数据同时考虑图像表示和散射特性两个重要因素, 利用物理散射特性约束视觉特征的学习, 在极少的标记样本情况下也能取得较好的分类效果。与第四章所提的基于时频特征的分类方法不同的是, 本章提出的框架可以统一地应用到所有极化方式的 SAR 图像中, 并且以无监督方法为主, 混合了深度神经网络和概率生成模型, 适用于在大的 SAR 场景下进行尺度灵活的精细地物分类。

为了实现这个过程, 首先本章从极化角度完善了上一章提出的子带散射特性分类方法, 即提出极化扩展的子带散射特性, 以补充应用到所有极化方式的 SAR 图像上; 其次, 用隐含狄利克雷分布模型对给定大小的切片生成不同散射主题的组合, 作为 SAR 图像一个重要的抽象属性; 假设部分与散射相关的抽象属性同时也能由图像的空间信息来表达, 设计一个软约束函数用于限制空间特征表示, 由此形成无监督的训练过程; 最后依靠少量的标记样本完成分类过程。

本章内容安排如下, 第二节详细介绍了散射特性约束下的 SAR 图像分类框架, 第三节给出了两个实际任务在所提框架下的应用效果, 第四节为本章小结。

6.2 散射特性约束下的分类框架

本节提出散射特性约束下的 SAR 图像分类框架, 总体来说一共分为五个模块: 空间特征提取、散射特性提取、散射主题生成、软约束学习、以及分类器学习, 如图6.2所示。在场景的语义类别中, 不同观测范围会产生不同的场景语义,

例如“机场”的语义需要在至少几百平方米的视角下进行理解，而“飞机”的语义理解所需的场景范围就相对而言较小。因此这里提出针对一个大型的 SAR 场景，对任意大小的 SAR 图像切片进行场景语义上的分类。主要的思路如下：

- 1) 给定大小的 SAR 幅度图像切片可通过空间特征提取网络得到视觉上的特征表示；
- 2) 同时对大型的 SAR 场景，通过散射特性提取可以得到逐目标的散射特性分类结果，如此一来，给定大小的 SAR 图像切片均有一个目标散射特性类别的切片相对应；
- 3) 通过散射主题生成模型，对每个散射特性切片生成所属的散射主题组合，作为对应 SAR 图像切片的固有散射属性；
- 4) 生成的散射属性和散射特性的类别相比具有更抽象的意义，与语义的抽象程度更加类似，因此假设一部分的散射属性也可以通过空间信息在另一个线性映射的空间中进行表示，由此设置软约束进行学习；
- 5) 最后利用少量的标记样本进行分类器的学习。

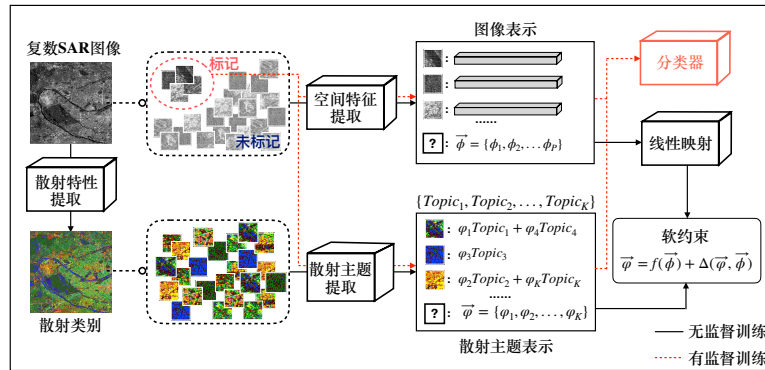


图 6.2 散射特性约束下的 SAR 图像分类框架

Figure 6.2 The SAR image classification framework under the constraint of scattering characteristics.

图6.2中所标注的黑色流程为无监督过程，红色虚线为监督学习过程，可以看到整个框架以无监督学习为主，不需要大量的标记样本来训练深度神经网络。所提框架的每个组成部分都具有一般性和可扩展性，为 SAR 图像分类提供了一种结合视觉和散射信息的新视角。例如空间特征提取模块可以通过第三章提出的 SAR 图像空间特征预训练模型的迁移学习得到，网络主干结构为 ResNet-18。下面将详细介绍其他几个组成部分可行的实现方法。

6.2.1 极化扩展的子带散射特性提取

散射特性提取模块可以采用常见的极化分析方法，得到如图5.8(a)所示的结果，当缺乏极化信息时，可由第五章提出的子带散射特性提取方法进行无监督学习得到。本节在第五章的基础上提出一种新的散射特性提取方案，将极化信息和子带散射信息相结合，学习极化扩展的子带散射特性，使此框架在极化 SAR 和非极化 SAR 的应用中具有统一性。

这个思路来源于Ferro-Famil 等 (2003) 提出极化性质会随方位角发生改变，而这种改变同时也反映了地面目标随方位角变化呈现出的非平稳性特征，因此本节提出对双极化和全极化的 SAR 数据结合子带散射图和极化特征，形成极化子带散射图进行极化子带散射分类，得到更丰富的散射类别结果。

对单通道的 SAR 图像中位于点 $l = (x_0, y_0)$ 的目标，和以此为中心的小邻域 $S(l)$ ，令距离向和方位向上分别以中心频率为 f_{ri} , f_{aj} 划分 R 个频带，即 $i, j = 1, 2, \dots, R$ ，对应的频域滤波器记为 $w(f_{ij})$ ，则点 l 处的目标在所有二维子带的散射功率变化可写为

$$Span_{TF} = [|r(f_{ij})|^2]_{R \times R}, \quad \dots (6.1)$$

其中

$$r(f_{ij}) = \text{IFFT}\{\text{FFT}[S(l)w(f_{ij})]\}_l \quad \dots (6.2)$$

对于极化 SAR 图像，将每个通道的数据进行二维子带分解后得到时频极化矩阵 S_{TF} ，定义为

$$S_{TF} = \begin{bmatrix} S_{HH}(f_{ij}) & S_{HV}(f_{ij}) \\ S_{VH}(f_{ij}) & S_{VV}(f_{ij}) \end{bmatrix}_{R \times R} \quad \dots (6.3)$$

则全极化 SAR 的时频 Pauli 基矢量 \vec{k}_{ij} 可写为

$$\vec{k}_{ij} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH}(f_{ij}) + S_{VV}(f_{ij}) & S_{HH}(f_{ij}) - S_{VV}(f_{ij}) & 2S_{HV}(f_{ij}) \end{bmatrix}^T, \quad \dots (6.4)$$

类似地，不同双极化 SAR（包括同极化 HH/VV 和交叉极化 XX/XY）的时频 Pauli 基矢量可分别定义为

$$\vec{k}_{ij} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH}(f_{ij}) + S_{VV}(f_{ij}) & S_{HH}(f_{ij}) - S_{VV}(f_{ij}) \end{bmatrix}^T, \quad \dots (6.5)$$

$$\vec{k}_{ij} = \begin{bmatrix} S_{XX}(f_{ij}) & 2S_{XY}(f_{ij}) \end{bmatrix}^T. \quad \dots (6.6)$$

由此可以得到时频极化相干矩阵 \mathbf{T}_{TF} ，定义为 $[\mathbf{T}_{ij}]_{R \times R}$ ，其中

$$\mathbf{T}_{ij} = \langle \vec{\mathbf{k}}_{ij} \cdot \vec{\mathbf{k}}_{ij}^\dagger \rangle. \quad \dots (6.7)$$

本节采用 [Cloude 和 Pottier \(1997\)](#) 提出的 $H/\alpha/A$ 极化分解方法对时频极化相干矩阵进行矩阵分解，提取出三个典型的时频极化特征，分别是时频极化熵 $\mathbf{H}_{\text{TF}} = [\mathbf{H}_{ij}]_{R \times R}$ 、时频平均极化角 $\boldsymbol{\alpha}_{\text{TF}} = [\alpha_{ij}]_{R \times R}$ 、以及时频极化总功率 $\mathbf{Span}_{\text{TF}} = [\mathbf{Span}_{ij}]_{R \times R}$ 。时频极化总功率反映了所有极化通道二维子带散射功率之和的变化，例如在全极化 SAR 的情况下

$$\mathbf{Span}_{ij} = |\mathbf{S}_{\text{HH}}(f_{ij})|^2 + |\mathbf{S}_{\text{HV}}(f_{ij})|^2 + |\mathbf{S}_{\text{VH}}(f_{ij})|^2 + |\mathbf{S}_{\text{VV}}(f_{ij})|^2. \quad \dots (6.8)$$

对时频极化相干矩阵 $[\mathbf{T}_{ij}]_{R \times R}$ 进行特征值分解，得到时频特征值 $[\lambda_{ij}^{(m)}]_{R \times R}$ 和对应的时频特征向量 $[\mathbf{u}_{ij}^{(m)}]_{R \times R}$ 。用 $[\mathbf{P}_{ij}^{(m)}]_{R \times R}$ 来表示每个特征值的概率，即

$$\mathbf{P}_{ij}^{(m)} = \frac{\lambda_{ij}^{(m)}}{\sum_{m=1}^n \lambda_{ij}^{(m)}}, \quad \dots (6.9)$$

其中全极化情况下 $n = 3$ ，双极化情况下 $n = 2$ 。这样一来得到时频极化熵矩阵 $[\mathbf{H}_{ij}]_{R \times R}$

$$\mathbf{H}_{ij} = - \sum_{m=1}^n \mathbf{P}_{ij}^{(m)} \log \mathbf{P}_{ij}^{(m)} \quad \dots (6.10)$$

以及时频平均极化角 $[\alpha_{ij}]_{R \times R}$

$$\alpha_{ij} = \sum_{m=1}^n \mathbf{P}_{ij}^{(m)} \cos^{-1}(|\mathbf{u}_{ij}^{(m)}(1)|) \quad \dots (6.11)$$

为了获得统一的输入，利于 HDEC-TFA 方法进行大规模的无监督学习，这里规定单极化 SAR 的输入模式为 $[\mathbf{Span}_{\text{TF}}, \mathbf{Span}_{\text{TF}}, \mathbf{Span}_{\text{TF}}]_{3 \times R \times R}$ ，极化 SAR 的输入模式为 $[\mathbf{H}_{\text{TF}}, \boldsymbol{\alpha}_{\text{TF}}, \mathbf{Span}_{\text{TF}}]_{3 \times R \times R}$ 。图6.3展示了对同一个目标在全极化、双极化 (HH/VV) 和单极化 (HH) 情况下的极化扩展子带散射图，作为 HDEC 网络的输入信号进行学习。

6.2.2 散射属性提取

根据 SAR 场景的散射特性结果，本节提出采用隐含狄利克雷分布模型 (Latent Dirichlet Allocation, LDA) 对每个 SAR 图像切片生成散射主题的混合表示，类

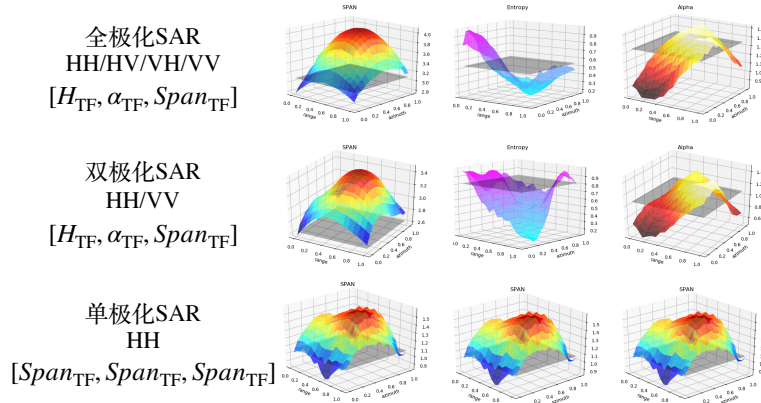


图 6.3 对不同极化方式的极化子带散射图输入。

Figure 6.3 The polarimetric sub-band scattering patterns for different polarimetric SAR data.

比于词袋模型 (Bag of Words, BoW) 用一组无序的单词 (words) 来表达一段文字, 在这里将这种用一组散射主题的组合来表达 SAR 图像切片属性的做法, 称为 BoST 模型 (Bag of Scattering Topics)。

LDA(Blei 等, 2003) 作为一种主题模型 (topic model) 最早提出来应用于文本挖掘, 本质上是一个三层的贝叶斯模型。它假设一篇文档 (document) 是由词 (word) 组成的集合, 词之间没有顺序关系, 每篇文档可以由多个主题 (topic) 来表示, 而文档中的每个单词都由主题以一定概率来生成。LDA 以概率分布的形式描述了文档-主题-词之间的关系, 不需要标注数据就能进行无监督的学习, 后来在图像领域也有相关的应用 (Rasiwasia 和 Vasconcelos, 2013; Bahmanyar 等, 2015)。

记大型 SAR 场景图像为 X , 对应的散射特性结果图记为 $\mathcal{L}(X)$, 其中共包括 N_s 种散射类别。根据 LDA 模型在图像领域的应用, 本节针对散射特性定义以下几个概念:

- 散射词向量 在散射特性结果图 $\mathcal{L}(X)$ 中随机选择一个 8×8 的小区域 \mathcal{L}_i , 将 \mathcal{L}_i 的归一化直方图系数定义为散射词向量, 即 $\vec{w}_i = \text{hist}(\mathcal{L}_i) = [h_1, h_2, \dots, h_{N_s}]$ 。

- 散射词典和散射词原子 在散射特性结果图 $\mathcal{L}(X)$ 中随机生成足够数量的散射词向量, 利用 k -means 算法 (Hartigan 和 Wong, 1979) 进行聚类, 得到 N_v 个聚类中心的集合作为散射词典, 即 $V = \{\vec{u}_1, \dots, \vec{u}_{N_v}\}$, 词典中的元素 \vec{u}_i 称为散射词原子。散射词原子即对应经典 LDA 算法中的“词”。

• 散射文档 根据待分类切片的大小，选取对应大小的散射特性结果图记为 $\mathcal{L}(x)$ 。以步幅为 4 在 $\mathcal{L}(x)$ 中依次选取大小为 8×8 的区域提取散射词向量 \vec{w}_d ，并利用学习的 k -means 模型找到其对应的散射词原子 \vec{u}_d ，最后统计 $\mathcal{L}(x)$ 中每个散射词原子出现的频率 r_d ，表示为散射文档，即 $d(x) = [r_1, r_2, \dots, r_{N_v}]$ ，对应经典 LDA 算法中的“文档”。

• 散射文档库 根据待分类切片的大小，在 $\mathcal{L}(X)$ 中随机选取数量为 N_d 个大小统一的散射结果图 $\mathcal{L}(x_i)$ 并计算散射文档 $d(x_i)$ ，组成散射文档的集合，即散射文档库 $D = \{d(x_1), d(x_2), \dots, d(x_{N_d})\}$ ，矩阵大小为 $N_v \times N_d$ 。散射文档库即为训练 LDA 主题模型时的数据，对应经典 LDA 算法中的“语料库”。

LDA 构建了一个“分布的分布”的三层生成式主题模型，生成一个文档的过程如下。LDA 假设主题和词的先验分布都是狄利克雷分布，首先选择一篇文档 d ，按照 LDA 模型的假设，从一个 K 维的狄利克雷分布 $Dir(\alpha)$ 中抽样生成文档 d 的主题分布 θ_d ，其中 K 是设置的主题个数；从主题的多项式分布 $Multinomial(\theta_d)$ 中抽样生成文档中第 n 个词的主题 z_{dn} ；由于词的先验分布服从 N_v 维的狄利克雷分布 $Dir(\beta)$ ，因此从 $Dir(\beta)$ 中可以抽样生成主题 z_{dn} 对应的词分布 ϕ_{dn} ，最终从词的多项式分布 $Multinomial(\phi_{dn})$ 中抽样生成文档 d 中的第 n 个词 u_{dn} 。

根据上述过程写出概率形式的表达。文档的主题分布 θ ，主题 z 和文档 d 的联合分布用概率形式可以写为

$$p(\theta, z, d | \alpha, \beta) = p(\theta | \alpha) \prod_{n=1}^{N_v} p(z_n | \theta) p(u_n | z_n, \beta) \quad \dots (6.12)$$

给定参数 α 和 β 的条件下，考虑 θ 和 z 所有的可能性，可以得到文档的边缘分布

$$p(d | \alpha, \beta) = \int p(\theta | \alpha) \left(\prod_{n=1}^{N_v} \sum_{z_n} p(z_n | \theta) p(u_n | z_n, \beta) \right) d\theta \quad \dots (6.13)$$

通过计算单个文档边缘分布的乘积，得到语料库的概率分布

$$p(D | \alpha, \beta) = \prod_{d=1}^{N_d} \int p(\theta_d | \alpha) \left(\prod_{n=1}^{N_v} \sum_{z_{dn}} p(z_{dn} | \theta_d) p(u_{dn} | z_{dn}, \beta) \right) d\theta_d \quad \dots (6.14)$$

LDA 模型的求解一般根据吉布斯采样或变分 EM 算法对上式参数进行估计。在本节的情况下，根据每张 SAR 图像切片的散射结果图 $\mathcal{L}(x)$ 得到散射文档 $d(x)$ ，然后用学习到的 LDA 模型生成散射主题分布，最终用 BoST 模型表示散射主题的组合，得到 SAR 图像的散射属性，整个过程记为

$$\varphi = \mathcal{G}(\mathcal{L}(x)) = \{\varphi_1, \varphi_2, \dots, \varphi_K\}, \quad \dots (6.15)$$

其中 φ_k 代表 LDA 模型预测的第 k 个散射主题的概率。

6.2.3 散射属性约束下的无监督学习

Bahmanyar 等 (2015) 比较了用词袋模型 (BoW) 和主题混合模型 (BoT) 来描述遥感类图像时的分类结果, 研究表明 BoT 提供了更简洁和抽象的语义级表示, 通过可视化 LDA 学习到的“主题”概念可以发现其与地物类别的联系。受相关研究的启发, 本节认为前述学习到的 SAR 图像切片散射主题组合反映了与散射特性相关的语义级信息, 这些语义级的信息与实际的地物类别具有一定的相关性, 那么同样地从视觉信息出发也能产生抽象的语义层级的理解, 因此本节提出重要假设: 将 $BoST$ 表示作为 SAR 图像切片固有的抽象属性, 其中一部分重要的属性也能通过视觉信息来表示。

在视觉层面上, 对 SAR 图像切片 x , 迁移第三章提出的 ResNet-18 幅度图像预训练模型得到空间视觉特征 $\phi = \mathcal{F}_{\text{img}}(x)$, 因为预训练模型 \mathcal{F}_{img} 具有较深的网络层数, 迁移的高层特征已经具有较抽象的表达。通过一个线性变换将图像空间特征表示映射到主题空间中, 记为 $f(\phi)$ 。定义下面的软约束目标函数

$$SoftLoss(\varphi, \phi) = \sum_{k \notin \mathbf{K}_{vis}} \varphi_k \cdot f(\phi) - \sum_{k \in \mathbf{K}_{vis}} \varphi_k \cdot f(\phi), \quad \dots (6.16)$$

其中 \mathbf{K}_{vis} 代表能被视觉信息表示的主题, 目标函数的优化使视觉信息与不在 \mathbf{K}_{vis} 中主题之间的关系最小化, 最大化视觉信息和 \mathbf{K}_{vis} 中主题的关系。在具体的实现过程中借鉴了多标签分类的损失函数计算方法, 将散射属性看作放松的伪标签, 采用 Softmax 激活线性映射后的空间特征, 计算其与散射属性之间的多标签分类损失函数如下

$$SoftLoss(\varphi, \phi) = - \sum_{k=1}^K \left(\varphi_k \log \frac{1}{1 + e^{-f(\phi_k)}} + (1 - \varphi_k) \log \frac{e^{-f(\phi_k)}}{1 + e^{-f(\phi_k)}} \right) \cdot \delta_k \quad \dots (6.17)$$

这里的 δ_k 设置为激活项, 选择 SAR 切片的散射属性中 $\varphi_k > 0.9$ 的主题, 并以 80%-90% 的概率激活, 使得 $\delta_k = 1$, 否则 $\delta_k = 0$ 。这个操作使视觉信息只去学习表示最重要的一部分散射主题, 保证了视觉信息和散射属性之间的软约束。

无监督学习通过优化目标函数 6.17 来完成, 在这个过程中, 训练数据不局限于数据集中带标记的 SAR 图像切片, 而是可以随机地在 SAR 场景中截取图像加入训练。学习过程中, 空间特征提取网络 \mathcal{F}_{img} 和线性映射 f 中的参数随着梯度下降进行更新。

6.2.4 分类器设计思路

本节根据带标记训练样本的规模提出了两个方向的分类器设计思路。当带标记样本较少但尚且可以支撑简单的神经网络训练的情况下，提出

(1) **ClsNet**: 在前述的网络结构基础上设计一个全连接层 g 作为到语义类别的映射，设计散射主题约束下的分类损失函数 L_{cls} ，定义为网络输出与类别间的交叉熵损失函数和软约束目标函数的加权求和形式：

$$L_{cls} = L_{CE}(g[f(\phi)], y) + \beta \cdot SoftLoss(\varphi, \phi), \quad \dots (6.18)$$

其中 y 为标记 SAR 图像的语义类别标签， β 控制软约束在总损失函数中所占的比例，此目标函数的优化可以通过梯度下降法完成，由此组成一个端到端的学习系统。只需用标记样本训练全连接层 g 中的参数，视情况微调前述神经网络中的参数，即可完成分类任务。

当带标记样本及其匮乏，例如每类只有几个可供训练使用的时候，本节给出以下建议来实现分类任务：

(2) **SiaNet**: 采用Koch等(2015)提出的孪生网络思路，设置两条相同的包括 \mathcal{F}_{img} 和 f 的支路，共享上述网络的参数，同时输入两个不同 SAR 图像切片和对应的散射分类结果图。一方面用对比损失函数评判两个样本之间特征 ϕ 的相似度，另一方面依旧将 $SoftLoss(\varphi, \phi)$ 作为软约束加入到训练过程中。将损失函数记为

$$L_{sia} = y \cdot |d(\phi_1, \phi_2)|^2 + (1-y) \cdot \max[0, m - d(\phi_1, \phi_2)] + \beta \cdot SoftLoss(\varphi, \phi) \quad \dots (6.19)$$

其中 d 表示两个特征之间的欧氏距离， m 为设置的边界，即不同类别的两个样本特征至少存在的欧氏距离，规定当输入的两个训练样本是同一类别时 $y = 1$ ，反之 $y = 0$ 。优化过程可同时微调 \mathcal{F}_{img} 和 f 中的参数，由于任意两个样本可组成一个输入对，这种方法在类别数较多但每类标记样本很少的情况下比较实用。测试时通过比较测试样本和训练集中的数据，找到最相似的样本即可得到预测结果。

(3) **SVM**: 将无监督学习得到的特征直接作为支撑向量机 (SVM) 的输入，用标记样本训练分类器，再对测试样本进行预测。

(4) **CosDist**: 用特征 ϕ 的余弦相似度衡量不同样本之间的距离，计算待测样本和每个类别所有标记样本的特征之间的平均余弦距离，找出最相似的类别作

为预测值，不需要额外的训练即可得到预测结果。两个不同特征间的余弦距离定义为

$$CosDist(\phi_1, \phi_2) = 1 - \frac{\phi_1 \cdot \phi_2}{\|\phi_1\|_2 \|\phi_2\|_2} \quad \dots (6.20)$$

6.3 实际应用

6.3.1 高分三号 SAR 图像城区精细地物分类

本节介绍的高分三号 SAR 图像城区精细地物分类任务的数据采用第5.4节介绍的巴黎地区全极化 SAR 图像，如图5.6(a) 所示。

为了验证第6.2.1节提出的极化扩展子带散射特性提取方法，本节首先从 SAR 场景中截取了大小为 512×512 像素的区域，分别选取单极化 HH 通道，双极化 HH/VV 通道，以及全极化四通道的 SAR 图像数据进行实验，利用第五章介绍的无监督学习方法得到极化扩展子带散射特性分类结果图，如图6.4所示。从图中可以明显分辨出加入极化信息对子带散射图进行扩展之后，算法对与自然地物（例如河流土地以及自然植被）相关的散射特性区分得更好，而人造目标相关的散射特性比较稳定，总的来说极化信息的辅助提升了模型对 SAR 图像散射特性的认知能力。

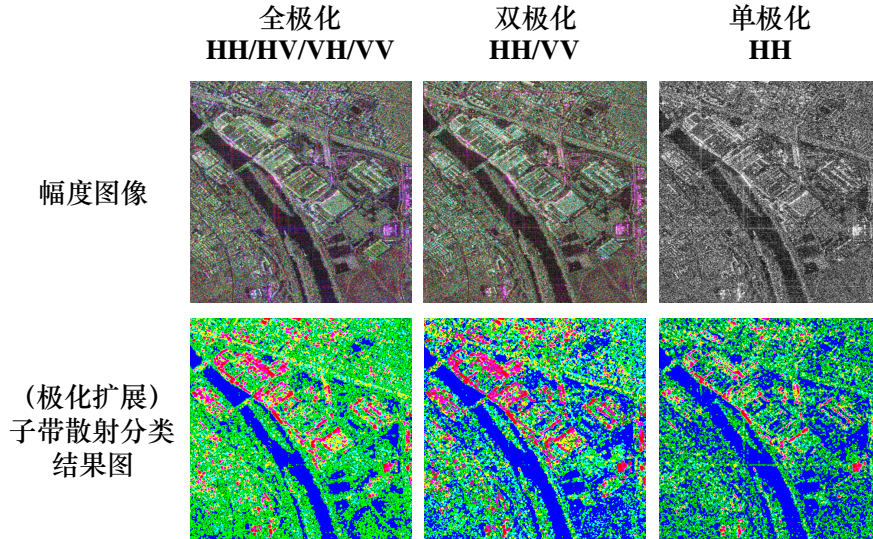


图 6.4 不同极化方式下的（极化扩展）子带散射特性分类结果示意图

Figure 6.4 The scattering pattern classification results for different polarimetric SAR data.

实验主要目的是对巴黎地区的城市地物进行精细化的语义分类，在这里将切片大小定为 64×64，图6.5中一共标记了 13 类地物类别共形成 233 个切片，类

别较为精细，包括桥梁、铁路等人造目标，体育馆、居民区、摩天大楼等建筑相关类别，十字路口、广场等与人类活动场所相关的类别，以及草地、河流、池塘等自然类目标。由于每个城市的特殊性，有的地物类别相对来说比较稀少，在实验过程中仅采用 4 个样本作为训练集，剩下的作为测试样本，如图 6.5 所示。

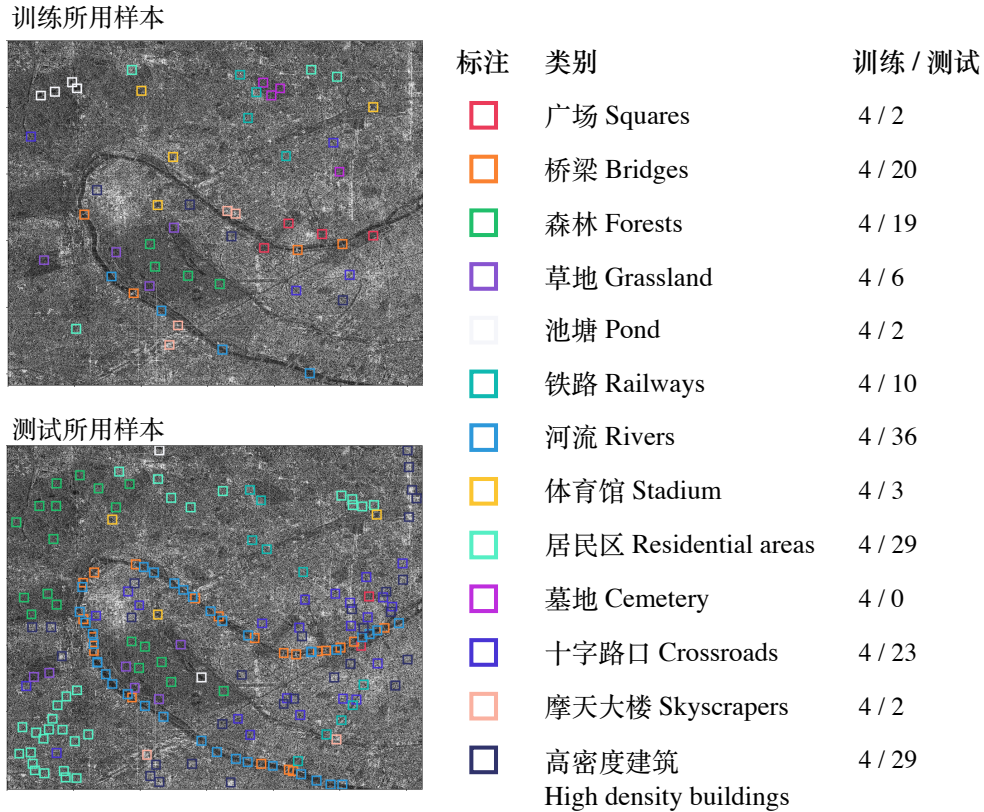


图 6.5 训练样本和测试样本示意图

Figure 6.5 The description of training and test data.

在这里选取 HH 通道的单极化 SAR 图像得到的子带散射特性结果，即图 5.8(a) 所示结果，来进行散射主题的提取。将 8×8 的小区域的散射特性直方图作为散射词， 64×64 大小的区域作为散射文档提取散射词向量表示，应用 LDA 算法得到每个 SAR 图像切片对应的散射主题组合表示。图 6.6 用白色正方形标记出四个代表性的散射主题为主的区域，从可视化的角度可以大致找出一些地物语义，使之与这些区域相关联。例如以散射主题 4 为主的区域与树林和墓地相关联，以散射主题 45 为主的区域与河流和草地有关联，以散射主题 24 为主的区域和高密度建筑较为相关。由此可以得出，更抽象的散射主题提取使得在场景中分布较为离散的物理散射特性更接近地物的语义层级。

考虑到本实验中的训练数据每类只有 4 个样本，因此这里采用第 6.2.4 节介

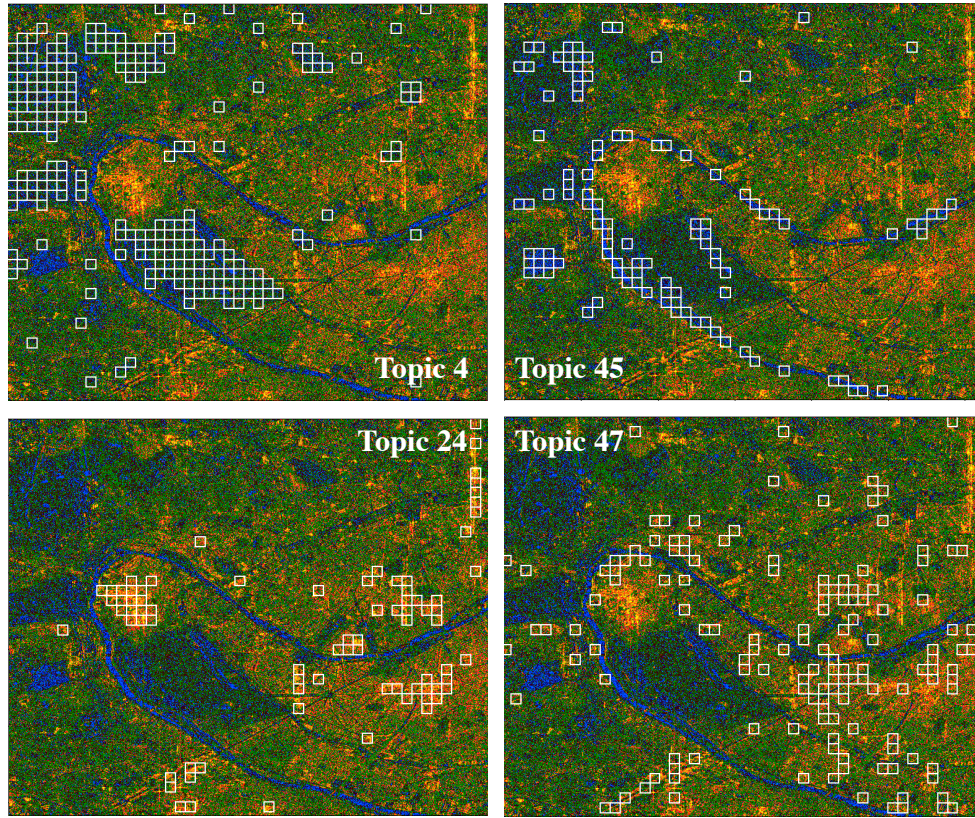


图 6.6 散射主题提取结果示意图

Figure 6.6 The scattering topic extraction.

绍的 (2)SiaNet、(3)SVM 和 (4)CosDist，三种分类方案。为了验证本方法的有效性，实验对比了仅采用预训练网络提取空间特征，和本章提出的加入散射特性之后得到的特征的分类结果，记录了 Top-1 和 Top-2 的测试集准确率，如表6.1所示。结果表明通过学习迁移的空间特征和散射特性之间的制约关系，提高了特征对 SAR 图像的代表能力，使其既具有空间纹理信息，也具有物理散射信息。在 SVM 分类器的情况下，Top-1 和 Top-2 测试准确率分别提高了 16.58% 和 23.2%。即使是最简单的不用额外训练的分类方法 CosDist，Top-1 测试准确率也提高了 15.6%，达到 56.35%，充分说明散射特性约束下的特征表示学习加大了特征的区分度，有效缓解了只学习空间信息时空间特征相似但实质上地物类别不一致的问题。从结果中可以看出，即使每类只用了 4 个标记样本进行训练，在 Top-2 的预测中也能达到 83.42% 的准确率，在实际应用中可以为人类提供非常有用的参考信息。

表 6.1 Top-1 和 Top-2 测试准确率

Table 6.1 The Top-1 and Top-2 test accuracies.

监督方法	CosDist	SiaNet	SVM
未加入散射特性 (Top-1)	0.4199	0.4501	0.4254
加入散射特性 (Top-1)	0.5635	0.5304	0.5912
未加入散射特性 (Top-2)	0.6519	0.6200	0.6022
加入散射特性 (Top-2)	0.8011	0.7459	0.8342

6.3.2 哨兵一号双极化 SAR 图像海冰分类

本节介绍本章所提方案在哨兵一号双极化 SAR 图像数据海冰分类中的应用。海冰监测对全球气候、大气环流以及海洋水循环等都具有重要的意义，相比起地物分类问题，海冰分类更加困难。冰川位置和大小时刻变化，受天气季节等因素的影响较大，而人类实地观测较为困难，类别标注十分受限。而且海冰根据状态分为多种类别，在某些观测条件下不同海冰类别的 SAR 的后向散射系数比较相近，难以分辨。

本次实验采用的海冰数据位于格陵兰岛东北部海岸地区，大约在北纬 79° 东经 16° 的位置，常年有海冰覆盖。Sentinal-1 卫星提供了宽幅模式 (IW) 下的双极化 SAR 图像数据，极化方式为 HH/HV，图6.7中列出了双极化复数 SAR 图像数据的幅度可视化图，以及处理后的地距图像示意图，其中地距分辨率约为 20 米，像素间隔约为 10 米，其入射角约为 39° 。

将地距幅度 SAR 图像以 256×256 的大小进行无重叠的切片，根据加拿大极冰局 (Canadian Ice Service, CIS) 提供的冰况图进行标注，具体标记的海冰类别和数量如表6.2所示，一共包括七个类别。其中区分较困难的三类，初期冰大约厚 10 到 30 厘米；一年冰是由初期冰发展成的厚冰，厚度在 30 厘米至 3 米之间，时间不超过一个冬季；老年冰是至少经过一个夏季未融化的冰，表面比一年冰更平滑。本节实验分别选取了 10% 和 5% 的数据作为训练样本，剩下的作为测试集，具体每类划分的数量详情见表6.2。

第6.2节提到，本章提出的分类框架其中的散射特性提取模块可以灵活性地采用第6.2.1节提出的极化扩展的子带散射特性提取方法，或者就是极化分解所

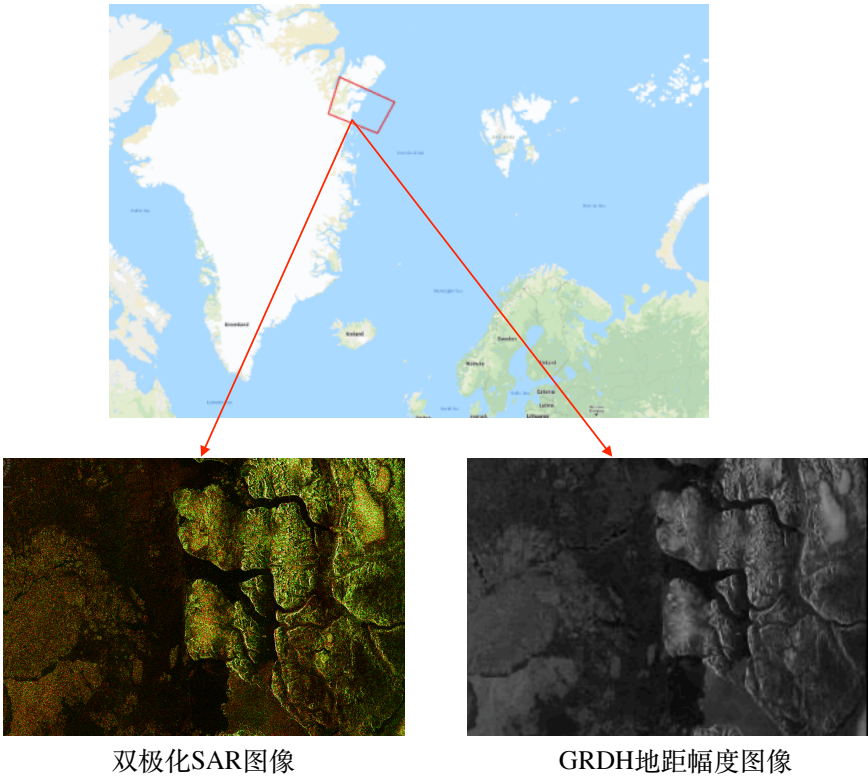


图 6.7 格陵兰岛所选实验数据示意图

Figure 6.7 The selected area of Greenland for experiments.

表 6.2 哨兵一号海冰分类数据详情

Table 6.2 The data description of sea-ice classification for Sentinal-1 Dual-Pol SAR data.

类别	水域	浮冰	冰山	冰河	老年冰	一年冰	初期冰
	Water	Floating Ice	Iceberg	Glaciers	Old Ice	First Year Ice	Young Ice
数目	506	141	230	101	302	204	316
训练集 (10%)	50	14	23	10	30	20	31
测试集 (90%)	456	127	207	91	272	184	285
训练集 (5%)	25	7	12	5	15	10	15
测试集 (95%)	481	134	218	96	287	194	301

得到的散射特性结果图。因此这里直接将 Sentinel-1 的双极化 SAR 数据的 H/α -Wishart 分类 (Cloude 和 Pottier, 1997) 的结果作为散射信息的来源，根据 Cloude 和 Pottier (1997) 提到的极化分解方法，最终的散射特性根据极化熵和散射角的关系被分成九个空间，其中八个是有效空间，代表明确的物理散射特性，例如二面角散射器、偶极子、随机表面散射等。 H/α -Wishart 分类结果如图6.8所示。

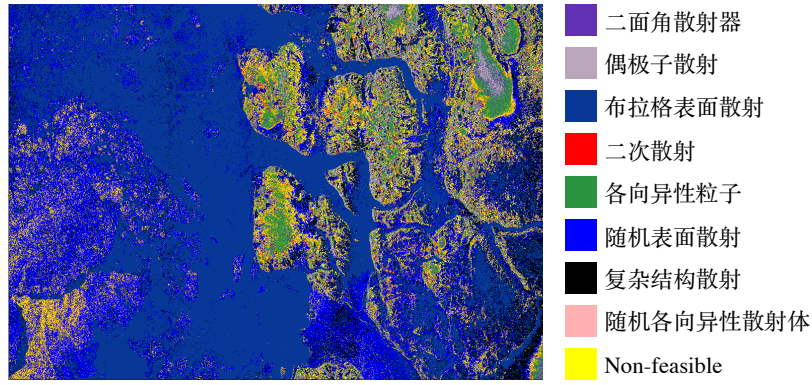


图 6.8 H/α-Wishart(Cloude 和 Pottier, 1997) 分类结果

Figure 6.8 The H/α-Wishart classification result.

经过处理之后，得到切片 SAR 图像与对应的散射特性结果图示例如图6.9所示，然后应用 LDA 方法生成每个散射特性结果图的散射主题组合，作为对应 SAR 图像切片的固有散射属性，约束神经网络的训练。

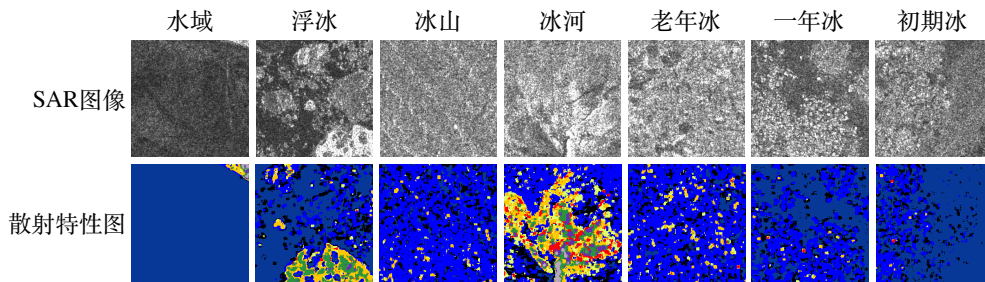


图 6.9 SAR 图像切片和对应散射特性图示例

Figure 6.9 The examples of SAR patches and the corresponding scattering results.

根据表6.2中给出的采用 5% 和 10% 的样本作为训练集时每类的数量，本节实验采取第6.2.4节中介绍的 (1)ClsNet 方案进行分类，记为 Scat-ClsNet。另外本节设置了其他几个对比实验，分别是只利用视觉信息迁移学习 SAR 图像预训练网络的 CNN，Dumitru 等 (2019) 所提 SAR 图像海冰分类算法 CAE-SVM 和 CAE-kNN，以及传统的特征提取分类算法 Gabor-SVM 和 Gabor-kNN，结果如表6.3所示。

表 6.3 哨兵一号 SAR 图像海冰分类实验结果

Table 6.3 The result of sea-ice classification for Sentinel-1 SAR images.

方法	Scat-ClsNet	CNN	CAE-SVM	CAE-kNN	Gabor-SVM	Gabor-kNN
10% Training	82.43%	80.97%	81.02%	80.13%	62.39%	64.24%
5% Training	82.02%	78.81%	78.14%	79.08%	56.05%	58.67%

6.4 本章小结

本章提出一种结合视觉特征和散射特性的 SAR 图像分类框架，融合深度学习方法和概率生成模型，在极少标注 SAR 图像样本的条件下实现分类任务。首先在第五章的基础上提出极化扩展的子带散射特性提取方法，为所有极化方式的 SAR 图像提供了统一的散射特性学习输入。对于不管是极化分析得到的极化散射特性分类结果，还是极化扩展的无监督子带散射特性分类结果，都可以应用主题模型学习到 SAR 图像在抽象语义层面的散射主题混合表示。将其作为 SAR 图像固有的属性，假设其中一部分属性也能通过视觉特征来表示，设计目标函数作为散射属性与图像视觉特征之间的软约束，对网络进行无监督的学习。最后提出几种分类器的设计方案，借助极少的标记样本完成分类任务。本章提出的框架具有一般性，适用于不同极化方式的大型 SAR 场景的理解；具有可扩展性，各个模块可以灵活设计具体实现方式；以及具有可解释性，揭示了散射特性在理解 SAR 图像过程中起到的重要辅助性作用。

本章的主要工作被 2020 年 IGARSS 国际会议接收为口头报告：

- HUANG Z, DATCU M, PAN Z, LEI B. A hybrid and explainable deep learning framework for SAR images, **IEEE International Geoscience and Remote Sensing Symposium 2020**, 口头报告

第7章 总结与展望

7.1 工作总结

SAR 图像自动分类是 SAR 图像智能解译领域中一个基础且重要的研究内容，包括目标识别、地物场景分类等任务。一方面作为“图像”，深度卷积神经网络被广泛应用到 SAR 图像分类中；而另一方面图像来自主动微波传感器“合成孔径雷达”，其应用与普通的光学图像相比又具有特殊性。目前深度卷积神经网络在 SAR 图像自动分类中的应用存在以下几个瓶颈：首先，由于 SAR 图像的目视解译需要专业的判读知识，因此对 SAR 图像进行大规模的人工标注较为困难，有限的标记样本很容易使深度神经网络在训练过程中出现过拟合；其次，SAR 图像在视觉上与光学图像存在很大的不同，但人眼同时也能从 SAR 图像中获取视觉上的信息，因此在深度学习方法的应用中既不能简单地将二者混为一谈，也不能完全地割裂二者的关系，关于这一点需要进行更深度的探讨；最后，SAR 图像反映的物理散射特性在 SAR 图像的理解上起着很大的辅助性作用，在缺乏足够坚实的物理模型和理论依据的支撑下，对于利用深度学习去发现潜在的散射特性，以及在 SAR 图像分类任务中同时从视觉层面和物理层面开展深度学习方法研究，目前国内外已发表文献涉及的还比较少。

本文面向 SAR 图像分类任务，立足于深度学习，特别是深度卷积神经网络，针对上述的几个问题开展研究，主要取得了以下创新性成果：

第二章从视觉层面出发进行 SAR 图像分类的空间特征迁移学习研究，针对 SAR 图像不同于光学图像的视觉特征，探究迁移学习在 SAR 目标识别应用中的特殊性。首先提出一种基于大规模无标签 SAR 图像的迁移学习方法，在逐渐减少训练样本的情况下能较好地保持模型的泛化能力，由于无标签 SAR 图像获取门槛低，此方法能得到广泛的应用。除此之外，第二章从数据、任务和网络三个方面进行讨论，提出在 SAR 图像中更有效地进行迁移学习可以采取的策略，并详细讨论了光学预训练模型在 SAR 图像的迁移中所存在的问题和解决方法，提出域适应传递式迁移学习，同时改善中低层和高层特征在 SAR 图像分类任务中的泛化能力，对面向 SAR 图像的迁移学习研究具有重要的参考价值。

第三章在第二章讨论结果的基础上，利用 DLR 提供的 TerraSAR-X 大规模

地物分类数据集，提出采用传递式迁移学习方法基于 ResNet-18 网络架构训练能广泛应用到 SAR 图像空间特征提取上的深度神经网络预训练模型。实验证明将其迁移到不同的 SAR 图像分类任务中均表现优异，对不同分辨率和不同传感器的 SAR 图像都具有良好的空间特征提取能力。本章开源了 SAR 图像预训练模型，为后续研究提供了快速的 SAR 图像空间特征提取渠道，并服务于社区各类与 SAR 图像相关的应用。

第四章创造性地提出了一种基于时频信号特征的 SAR 图像分类框架。以单极化复数 SAR 图像作为研究对象，以二维连续子带分解为手段，在距离向和方位向的连续子带上依次获得低分辨率子带 SAR 图像，产生同时包含二维空间信息和距离方位向两个维度频域信息的四维时频信号。然后通过设计神经网络模型从复数信号中提取频域特征，同时融合迁移学习的幅度图像空间纹理特征作为 SAR 图像的代表进行分类。这种方法与纯数据驱动的复数神经网络相比，能更好地应用已知的 SAR 图像的性质，减轻了神经网络学习的负担，并极大改善了传统的基于空间特征的深度卷积神经网络模型对于视觉上相似的地物类别难以区分的问题。

第五章首次提出一种自动提取单极化 SAR 图像物理散射特性的无监督深度学习方法。针对复数 SAR 图像中的每个位置的目标，基于时频分析理论提取子带散射图，反映了目标在不同方位角和不同电磁波频段上散射强度的变化。设计多层级的深度嵌入聚类方法自动学习子带散射图的特征并进行分类，得到最终的子带散射特性分类结果图。与基于极化散射矩阵的极化分类过程类似，整个过程在无监督学习框架下完成，但不同的是极化散射类型具有明确的物理含义，可以用极化物理散射模型进行解释，而子带散射类型在单纯数据驱动的神经网络训练下获得，物理层面上的可解释性较差。但通过基于信息论的定量化评估方法，可以挖掘出子带散射类型和极化散射类别之间的潜在关系，从侧面说明了提取出的单极化 SAR 图像子带散射类型反映了目标的物理特性。

第六章首次提出散射特性约束下的 SAR 图像分类新框架，结合深度神经网络和概率生成模型等，无论是否包含极化信息都从物理层面和视觉层面对 SAR 图像分类进行研究。首先在极化层面上扩展了第五章提出的子带散射类别分类方法，将算法应用于所有极化方式的复数 SAR 图像中，得到极化扩展的子带散射类别。针对需要分类的 SAR 图像大小，利用概率生成模型得到每个 SAR 图像

切片的散射主题分布,作为 SAR 图像潜在的虚拟属性,在散射属性的约束下改善 SAR 图像空间特征的表达,最终用极少数的标记样本实现分类。本章提出的新型框架具有灵活的可扩展性,例如在散射特性提取时也可以用极化分析的方法来实现,并可以根据标记样本的多少来选择最后监督分类的方法。实验将其应用到城市精细化分类和海冰分类两种任务中都取得了不错的效果。

7.2 研究展望

本文针对深度学习方法在 SAR 图像的分类任务中应用的几点局限性开展了研究,总体来说取得了一定的阶段性成果。但在目前研究成果的基础上,仍然存在许多问题需要进一步的深入研究。

SAR 图像专用的深度神经网络模型构建 本文在第三章提出了一个针对 SAR 幅度图像的迁移能力较强的预训练网络模型,随着 SAR 图像数据和解译任务的多样性逐渐丰富,对 SAR 图像专用预训练模型的需求会逐步提升。利用更丰富的非公开以及公开的 SAR 图像标记样本打造一系列不同网络骨架结构的预训练模型,将有利于社区更高效地进行基于深度学习方法的 SAR 图像解译研究。

小样本条件下的 SAR 图像分类问题 本文在第二章提出迁移学习在 SAR 图像应用中的特殊性策略,能缓解一部分缺少训练样本的问题,但对真正意义上的小样本学习还具有较大的差距。在有限标记 SAR 样本的条件下,重要的是能否尽可能地从大量的没有语义标记的样本中获取更多的知识。SAR 图像中蕴含的物理散射特性为这个思路提供了可能性,本文在第六章初步构建了一个同时从视觉和物理层面进行智能认知的框架,只用了极少数的标记样本来进行实验,但总体来说还不够成熟,未来将会延续这个思路进行研究。

可解释的基于深度学习的散射特性提取 本文在第五章提出了无监督的子带散射特性提取方法,但目前各种子带散射类别不像极化分析的结果一样具有很强的可解释性,而且一定程度上存在数据特殊性。子带散射图的特性是否可以用较稳定的特征空间来表示,并给出划分特征空间中类别边界的条件,从特征表示的角度去解释学习到的散射特性,是今后需要解决的问题。

定量化的认知结果表示 目前在 SAR 图像的智能解译研究中,深度学习并不能完全代替人类的认知,我们希望深度学习能给出定量化的认知结果辅助人类快速理解大规模的 SAR 图像。例如第六章介绍的实例中,仅依靠极少的标记数

据，深度网络能提供准确率较高的前两个最有可能的标签，但这远不够提供量化的参考。今后可以深入探究如何在深度学习中提出更多细致的量化指标，例如预测的置信度，视觉层面或物理层面对结果的贡献程度等。

参考文献

- 周雨, 王海鹏, 陈思吉吉. 基于数值散射模拟与模型匹配的 SAR 自动目标识别研究[J/OL]. 雷达学报, 2015, 4(6):666-673. <http://radars.ie.ac.cn/article/doi/10.12000/JR15080>.
- 文贡坚, 朱国强, 殷红成, 等. 基于三维电磁散射参数化模型的 SAR 目标识别方法[J/OL]. 雷达学报, 2017, 6(2):115-135. <http://radars.ie.ac.cn/article/doi/10.12000/JR17034>.
- 潘宗序, 安全智, 张冰尘. 基于深度学习的雷达图像目标识别研究进展[J/OL]. 中国科学: 信息科学, 2019, 49(12):1626-1639. DOI: <https://doi.org/10.1360/SSI-2019-0093>.
- AN Q, PAN Z, YOU H. Ship detection in Gaofen-3 SAR images based on sea clutter distribution analysis and deep convolutional neural network[J/OL]. Sensors, 2018, 18(2):334. <https://www.mdpi.com/1424-8220/18/2/334>. DOI: [10.3390/s18020334](https://doi.org/10.3390/s18020334).
- ANAGNOSTOPOULOS G C. SVM-based target recognition from synthetic aperture radar images using target region outline descriptors[J]. Nonlinear Analysis, Theory, Methods and Applications, 2009, 71(12):2934-2939.
- AZIZPOUR H, RAZAVIAN A S, SULLIVAN J, et al. From generic to specific deep representations for visual recognition[C/OL]//IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). 2015: 36-45. DOI: [10.1109/CVPRW.2015.7301270](https://doi.org/10.1109/CVPRW.2015.7301270).
- BAHMANYAR R, CUI S, DATCU M. A comparative study of bag-of-words and bag-of-topics models of EO image patches[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(6): 1357-1361.
- BAO X, PAN Z, LIU L, et al. SAR image simulation by generative adversarial networks[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2019: 9995-9998.
- BERRADA L, ZISSERMAN A, KUMAR M P. Smooth loss functions for deep top-k classification [C]//International Conference on Learning Representations (ICLR). 2018.
- BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation[J]. Journal of Machine Learning Research, 2003, 3(4-5):993-1022.
- BORGWARDT K M, GRETTON A, RASCH M J, et al. Integrating structured biological data by kernel maximum mean discrepancy[J]. Bioinformatics, 2006, 22(14):e49-e57.
- BOVENGA F, GIACOVAZZO V, REFICE A, et al. Interferometric multi-chromatic analysis of high resolution X-band data[C]//Fringe 2011 Workshop. 2011: 19-23.
- BOVENGA F, DERAUW D, RANA F M, et al. Multi-chromatic analysis of SAR images for coherent target detection[J/OL]. Remote Sensing, 2014, 6(9):8822-8843. <http://www.mdpi.com/2072-4292/6/9/8822>. DOI: [10.3390/rs6098822](https://doi.org/10.3390/rs6098822).

- CHEN S, SATO M. General polarimetric model-based decomposition for coherency matrix[C/OL]// IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2012: 99-102. DOI: [10.1109/IGARSS.2012.6351627](https://doi.org/10.1109/IGARSS.2012.6351627).
- CHEN S, TAO C. PolSAR image classification using polarimetric-feature-driven deep convolutional neural network[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(4):627-631. DOI: [10.1109/LGRS.2018.2799877](https://doi.org/10.1109/LGRS.2018.2799877).
- CHEN S, WANG H, XU F, et al. Target classification using the deep convolutional networks for SAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(8):4806-4817.
- CHEN Z, ZHANG T, OUYANG C. End-to-end airplane detection using transfer learning in remote sensing images[J/OL]. Remote Sensing, 2018, 10(1):139. <http://www.mdpi.com/2072-4292/10/1/139>. DOI: [10.3390/rs10010139](https://doi.org/10.3390/rs10010139).
- CHENG G, ZHOU P, HAN J. Learning rotation-invariant convolutional neural networks for object detection in VHR optical remote sensing images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(12):7405-7415.
- CHENG G, HAN J, LU X. Remote sensing image scene classification: Benchmark and state of the art[J/OL]. Proceedings of the IEEE, 2017, 105(10):1865-1883. DOI: [10.1109/JPROC.2017.2675998](https://doi.org/10.1109/JPROC.2017.2675998).
- CLOUDE S, POTTIER E. An entropy based classification scheme for land applications of polarimetric SAR[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 1997, 35(1):68-78. DOI: [10.1109/TGRS.2018.2866934](https://doi.org/10.1109/TGRS.2018.2866934).
- CUI Z, ZHANG M, CAO Z, et al. Image data augmentation for SAR sensor via generative adversarial nets[J]. IEEE Access, 2019, 7:42255-42268.
- Defense Advanced Research Projects Agency. The air force moving and stationary target recognition database[EB/OL]. <https://www.sdms.afrl.af.mil/datasets/mstar/>. Accessed 3 February, 2016.
- DENG J, DONG W, SOCHER R, et al. ImageNet: A large-scale hierarchical image database[C/OL]// IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2009: 248-255. DOI: [10.1109/CVPR.2009.5206848](https://doi.org/10.1109/CVPR.2009.5206848).
- DING J, CHEN B, LIU H, et al. Convolutional neural network with data augmentation for SAR target recognition[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(3):364-368.
- DONG G, WANG N, KUANG G. Sparse representation of monogenic signal: With application to target recognition in SAR images[J]. IEEE Signal Processing Letters, 2014, 21(8):952-956.
- DONG G, KUANG G, WANG N, et al. SAR target recognition via joint sparse representation of monogenic signal[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(7):3316-3328.
- DU P, SAMAT A, WASKE B, et al. Random forest and rotation forest for fully polarized SAR

- image classification using polarimetric and spatial features[J/OL]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 105:38 - 53. <http://www.sciencedirect.com/science/article/pii/S0924271615000611>. DOI: <https://doi.org/10.1016/j.isprsjprs.2015.03.002>.
- DUMITRU C O, ANDREI V, SCHWARZ G, et al. Machine learning for sea ice monitoring from satellites[J/OL]. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019, 4216:83-89. <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W16/83/2019/>. DOI: [10.5194/isprs-archives-XLII-2-W16-83-2019](https://doi.org/10.5194/isprs-archives-XLII-2-W16-83-2019).
- DUMITRU C O, SCHWARZ G, DATCU M. Land cover semantic annotation derived from high-resolution SAR images[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2016, 9(6):2215-2232.
- FERRO-FAMIL L, REIGBER A, POTTIER E, et al. Scene characterization using subaperture polarimetric SAR data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2003, 41(10): 2264-2276.
- FERRO-FAMIL L, REIGBER A, POTTIER E. Nonstationary natural media analysis from polarimetric SAR data using a two-dimensional time-frequency decomposition approach[J/OL]. *Canadian Journal of Remote Sensing*, 2005, 31(1):21-29. <https://doi.org/10.5589/m04-062>.
- FREEMAN A, DURDEN S L. A three-component scattering model for polarimetric SAR data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1998, 36(3):963-973.
- FU G, LIU C, ZHOU R, et al. Classification for high resolution remote sensing imagery using a fully convolutional network[J]. *Remote Sensing*, 2017, 9(5):498.
- GANIN Y, LEMPITSKY V. Unsupervised domain adaptation by backpropagation[C/OL]// *International Conference on Machine Learning (ICML): volume 37*. 2015: 1180-1189. <http://proceedings.mlr.press/v37/ganin15.html>.
- GENG J, DENG X, MA X, et al. Transfer learning for SAR image classification via deep joint distribution adaptation networks[J/OL]. *IEEE Transactions on Geoscience and Remote Sensing*, 2020:1-16. DOI: [10.1109/TGRS.2020.2964679](https://doi.org/10.1109/TGRS.2020.2964679).
- GENG J, FAN J, WANG H, et al. High-resolution SAR image classification via deep convolutional autoencoders[J]. *IEEE Geoscience and Remote Sensing Letters*, 2015, 12(11):2351-2355.
- GENG J, WANG H, FAN J, et al. Deep supervised and contractive neural network for SAR image classification[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(4):2442-2459.
- GENTINE P, PRITCHARD M, RASP S, et al. Could machine learning break the convection parameterization deadlock?[J/OL]. *Geophysical Research Letters*, 2018, 45(11):5742-5751. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018GL078202>.
- GRETTON A, SEJDINOVIC D, STRATHMANN H, et al. Optimal kernel choice for large-scale two-

- sample tests[C/OL]//Advances in Neural Information Processing Systems (NIPS). 2012: 1205-1213. <http://papers.nips.cc/paper/4727-optimal-kernel-choice-for-large-scale-two-sample-tests.pdf>.
- GUINVARC' H R, THIRION-LEFEVRE L. Cross-polarization amplitudes of obliquely orientated buildings with application to urban areas[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(11):1913-1917. DOI: [10.1109/LGRS.2017.2731122](https://doi.org/10.1109/LGRS.2017.2731122).
- GUO J, LEI B, DING C, et al. Synthetic aperture radar image synthesis by using generative adversarial nets[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(7):1111-1115.
- HARTIGAN J A, WONG M A. Algorithm as 136: A k-means clustering algorithm[J/OL]. Journal of the Royal Statistical Society. Series C (Applied Statistics), 1979, 28(1):100-108. <http://www.jstor.org/stable/2346830>.
- HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C/OL]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016: 770-778. DOI: [10.1109/CVPR.2016.90](https://doi.org/10.1109/CVPR.2016.90).
- HOU B, KOU H, JIAO L. Classification of polarimetric SAR images using multilayer autoencoders and superpixels[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016, 9(7):3072-3081.
- HOU X, AO W, XU F. End-to-end automatic ship detection and recognition in high-resolution Gaofen-3 spaceborne SAR images[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2019: 9486-9489.
- HU F, XIA G S, HU J, et al. Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery[J/OL]. Remote Sensing, 2015, 7(11): 14680-14707. <http://www.mdpi.com/2072-4292/7/11/14680>. DOI: [10.3390/rs71114680](https://doi.org/10.3390/rs71114680).
- HUANG G, LIU Z, WEINBERGER K Q. Densely connected convolutional networks[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017a: 2261-2269.
- HUANG L, LIU B, LI B, et al. OpenSARShip: A dataset dedicated to Sentinel-1 ship interpretation [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11(1):195-208.
- HUANG Z, PAN Z, LEI B. Transfer learning with deep convolutional neural network for SAR target classification with limited labeled data[J]. Remote Sensing, 2017, 9(9):907.
- HÜTT C, KOPPE W, MIAO Y, et al. Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images [J/OL]. Remote Sensing, 2016, 8(8):684. DOI: [10.3390/rs8080684](https://doi.org/10.3390/rs8080684).
- JIAO L, LIU F. Wishart deep stacking network for fast POLSAR image classification[J/OL]. IEEE Transactions on Image Processing, 2016, 25(7):3273-3286. DOI: [10.1109/TIP.2016.2567069](https://doi.org/10.1109/TIP.2016.2567069).

- JIAO X, KOVACS J M, SHANG J, et al. Object-oriented crop mapping and monitoring using multi-temporal polarimetric RADARSAT-2 data[J/OL]. ISPRS Journal of Photogrammetry and Remote Sensing, 2014, 96:38 - 46. <http://www.sciencedirect.com/science/article/pii/S0924271614001695>. DOI: <https://doi.org/10.1016/j.isprsjprs.2014.06.014>.
- KANG C, HE C. SAR image classification based on the multi-layer network and transfer learning of mid-level representations[C/OL]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2016: 1146-1149. DOI: [10.1109/IGARSS.2016.7729290](https://doi.org/10.1109/IGARSS.2016.7729290).
- KANG M, JI K, LENG X, et al. Contextual region-based convolutional neural network with multilayer fusion for SAR ship detection[J]. Remote Sensing, 2017, 9(8):860.
- KARPATNE A, WATKINS W, READ J S, et al. Physics-guided neural networks (PGNN): an application in lake temperature modeling[J/OL]. 2017. <http://arxiv.org/abs/1710.11431>.
- KOCH G, ZEMEL R, SALAKHUTDINOV R. Siamese neural networks for one-shot image recognition[C]//Proceedings of ICML deep learning workshop: volume 2. 2015.
- KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C/OL]//Advances in Neural Information Processing Systems (NIPS). 2012: 1097-1105. <http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf>.
- KUKAR M, KONONENKO I, et al. Cost-sensitive learning with neural networks[C]//European Conference on Artificial Intelligence (ECAI). 1998: 445-449.
- LEE J S, GRUNES M R, AINSWORTH T L, et al. Unsupervised classification using polarimetric decomposition and the complex wishart classifier[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 1999, 37(5):2249-2258. DOI: [10.1109/36.789621](https://doi.org/10.1109/36.789621).
- LEE J S, GRUNES M R, POTTIER E, et al. Unsupervised terrain classification preserving polarimetric scattering characteristics[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(4):722-731. DOI: [10.1109/TGRS.2003.819883](https://doi.org/10.1109/TGRS.2003.819883).
- LIN Z, JI K, KANG M, et al. Deep convolutional highway unit network for SAR target classification with limited labeled training data[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(7): 1091-1095.
- LIU G, SUN X, FU K, et al. Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior[J]. IEEE Geoscience and Remote Sensing Letters, 2013, 10(3):573-577.
- LIU H, LI S. Decision fusion of sparse representation and support vector machine for SAR image target recognition[J]. Neurocomputing, 2013, 113:97-104.
- LIU X, JIAO L, TANG X, et al. Polarimetric convolutional network for PolSAR image classification [J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(5):3040-3054. DOI: [10.1109/TGRS.2018.2879984](https://doi.org/10.1109/TGRS.2018.2879984).

- LONG M, CAO Y, WANG J, et al. Learning transferable features with deep adaptation networks [C/OL]//International Conference on Machine Learning (ICML): volume 37. 2015: 97-105. <http://proceedings.mlr.press/v37/long15.html>.
- LONG M, ZHU H, WANG J, et al. Deep transfer learning with joint adaptation networks[C/OL]// International Conference on Machine Learning (ICML): volume 70. 2017: 2208-2217. <http://proceedings.mlr.press/v70/long17a.html>.
- LU C, LI W. Ship classification in high-resolution SAR images via transfer learning with small training dataset[J]. Sensors, 2019, 19(1):63.
- LV Q, DOU Y, NIU X, et al. Urban land use and land cover classification using remotely sensed SAR data through deep belief networks[J]. Journal of Sensors, 2015, 2015.
- MALMGREN-HANSEN D, KUSK A, DALL J, et al. Improving SAR automatic target recognition models with transfer learning from simulated data[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(9):1484-1488.
- MARMANIS D, DATCU M, ESCH T, et al. Deep learning earth observation classification using ImageNet pretrained networks[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(1): 105-109.
- MARMANIS D, YAO W, ADAM F, et al. Artificial generation of big data for improving image classification: A generative adversarial network approach on SAR data[J/OL]. 2017. <https://arxiv.org/abs/1711.02010>.
- MEHRA R K, HUFF M, RAVICHANDRAN R B, et al. Nonparametric error estimation techniques applied to MSTAR data sets[C]//Algorithms for Synthetic Aperture Radar Imagery V: volume 3370. International Society for Optics and Photonics, 1998: 614-625.
- MINH D H T, IENCO D, GAETANO R, et al. Deep recurrent neural networks for winter vegetation quality mapping via multitemporal SAR sentinel-1[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15(3):464-468.
- MORGAN D A. Deep convolutional neural networks for ATR from SAR imagery[C/OL]// Proceedings of SPIE: volume 9475. 2015. <http://dx.doi.org/10.1117/12.2176558>.
- NDIKUMANA E, HO TONG MINH D, BAGHDADI N, et al. Deep recurrent neural network for agricultural classification using multitemporal SAR Sentinel-1 for Camargue, France[J]. Remote Sensing, 2018, 10(8):1217.
- NOVAK L M, OWIRKA G J, BROWER W S. Performance of 10- and 20-target MSE classifiers[J]. IEEE Transactions on Aerospace and Electronic Systems, 2000, 36(4):1279-1289.
- PAPSON S, NARAYANAN R M. Classification via the shadow region in SAR imagery[J]. IEEE Transactions on Aerospace and Electronic Systems, 2012, 48(2):969-980.

- PEI J, HUANG Y, HUO W, et al. SAR automatic target recognition based on multiview deep learning framework[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(4):2196-2210.
- PEI J, HUANG Y, SUN Z, et al. Multiview synthetic aperture radar automatic target recognition optimization: Modeling and implementation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(11):6425-6439.
- QIN F, GUO J, SUN W. Object-oriented ensemble classification for polarimetric SAR imagery using restricted boltzmann machines[J]. *Remote Sensing Letters*, 2017, 8(3):204-213.
- RASIWASIA N, VASCONCELOS N. Latent dirichlet allocation models for image classification[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(11):2665-2679.
- RATHA D, BHATTACHARYA A, FRERY A C. Unsupervised classification of PolSAR data using a scattering similarity measure derived from a geodesic distance[J/OL]. *IEEE Geoscience and Remote Sensing Letters*, 2018, 15(1):151-155. DOI: [10.1109/LGRS.2017.2778749](https://doi.org/10.1109/LGRS.2017.2778749).
- RATHA D, GAMBA P, BHATTACHARYA A, et al. Novel techniques for built-up area extraction from polarimetric SAR images[J/OL]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 17(1):177 - 181. DOI: [10.1109/LGRS.2019.2914913](https://doi.org/10.1109/LGRS.2019.2914913).
- REICHSTEIN M, CAMPS-VALLS G, STEVENS B, et al. Deep learning and process understanding for data-driven earth system science[J/OL]. *Nature*, 2019, 566(7743):195-204. <https://doi.org/10.1038/s41586-019-0912-1>.
- RENGA A, GRAZIANO M D, MOCCIA A. Segmentation of marine SAR images by sublook analysis and application to sea traffic monitoring[J/OL]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(3):1463-1477. DOI: [10.1109/TGRS.2018.2866934](https://doi.org/10.1109/TGRS.2018.2866934).
- RUAN H, ZHANG R, LI J, et al. SAR target classification based on multiscale sparse representation[C/OL]//ISPRS International Conference on Computer Vision in Remote Sensing (CVRS): volume 9901. 2016: 185-190. <https://doi.org/10.1117/12.2234825>.
- SAMAT A, GAMBA P, ABUDUWAILI J, et al. Geodesic flow kernel support vector machine for hyperspectral image classification by unsupervised subspace feature transfer[J]. *Remote Sensing*, 2016, 8(3):234.
- SATO R, YAMAGUCHI Y, YAMADA H. Polarimetric scattering feature estimation for accurate vegetation area classification[C/OL]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS): volume 3. 2009: III-888-III-891. DOI: [10.1109/IGARSS.2009.5417912](https://doi.org/10.1109/IGARSS.2009.5417912).
- SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[C/OL]//International Conference on Learning Representations (ICLR). 2015. <https://arxiv.org/abs/1409.1556>.
- SINGH J, DATCU M. SAR target analysis based on multiple-sublook decomposition: a visual exploration approach[J]. *IEEE Geoscience and Remote Sensing Letters*, 2012, 9(2):247-251.

- SONG Q, XU F, JIN Y. Radar image colorization: Converting single-polarization to fully polarimetric using deep neural networks[J]. IEEE Access, 2018, 6:1647-1661.
- SONG S, XU B, YANG J. SAR target recognition via supervised discriminative dictionary learning and sparse representation of the SAR-HOG feature[J]. Remote Sensing, 2016, 8(8):683.
- SPIGAI M, TISON C, SOUYRIS J C. Time-frequency analysis in high-resolution SAR imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 49(7):2699-2711.
- SUMBUL G, CHARFUELAN M, DEMIR B, et al. BigEarthNet: A large-scale benchmark archive for remote sensing image understanding[C]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2019: 5901-5904.
- SUN A Y, SCANLON B R, ZHANG Z, et al. Combining physically based modeling and deep learning for fusing grace satellite data: Can we learn from mismatch?[J/OL]. Water Resources Research, 2019, 55(2):1179-1195. <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2018WR023333>.
- SUN J, HAN B, YUAN X, et al. Gaofen-3 (GF-3) full (quad) polarimetric sample datasets[EB/OL]. <https://www.ietr.fr/GF3/>. Accessed 3 February, 2019.
- SUN J, YU W, DENG Y. The SAR payload design and performance for the GF-3 mission[J/OL]. Sensors, 2017, 17(10):2419. DOI: [10.3390/s17102419](https://doi.org/10.3390/s17102419).
- SUN Y, DU L, WANG Y, et al. SAR automatic target recognition based on dictionary learning and joint dynamic sparse representation[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12):1777-1781.
- SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions[C/OL]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015: 1-9. <http://dx.doi.org/10.1109/CVPR.2015.7298594>.
- TerraSAR-X Science Service System. TerraSAR-X basic product specification document, issue 1.9 [EB/OL]. <https://sss.terrasar-x.dlr.de/docs/TX-GS-DD-3302.pdf>. Accessed December 10, 2018.
- TONG X Y, XIA G S, LU Q, et al. Land-cover classification with high-resolution remote sensing images using transferable deep models[J/OL]. Remote Sensing of Environment, 2020, 237:111322. <http://www.sciencedirect.com/science/article/pii/S0034425719303414>. DOI: <https://doi.org/10.1016/j.rse.2019.111322>.
- TUPIN F, TISON C. Sub-aperture decomposition for SAR urban area analysis[J]. European Conference on Synthetic Aperture Radar (EUSAR), 2004:431-434.
- TZENG E, HOFFMAN J, ZHANG N, et al. Deep domain confusion: Maximizing for domain invariance[J/OL]. 2014. <https://arxiv.org/abs/1412.3474>.
- VALVERDE-ALBACETE F J, PELÁEZ-MORENO C. A framework for supervised classification

- performance analysis with information-theoretic methods[J/OL]. IEEE Transactions on Knowledge and Data Engineering, 2019. DOI: [10.1109/TKDE.2019.2915643](https://doi.org/10.1109/TKDE.2019.2915643).
- WANG K, ZHANG G, LENG Y, et al. Synthetic aperture radar image generation with deep generative models[J]. IEEE Geoscience and Remote Sensing Letters, 2018, 16(6):912 - 916.
- WANG L, SCOTT K A, XU L, et al. Sea ice concentration estimation during melt from dual-pol SAR scenes using deep convolutional neural networks: A case study[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(8):4524-4533.
- WANG S, QUAN D, LIANG X, et al. A deep learning framework for remote sensing image registration[J/OL]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 145:148-164. <http://www.sciencedirect.com/science/article/pii/S0924271617303891>. DOI: <https://doi.org/10.1016/j.isprsjprs.2017.12.012>.
- WANG Y, WANG C, ZHANG H, et al. Combining single shot multibox detector with transfer learning for ship detection using Chinese Gaofen-3 images[C/OL]//Progress in Electromagnetics Research Symposium (PIERS). 2017: 712-716. DOI: [10.1109/PIERS-FALL.2017.8293227](https://doi.org/10.1109/PIERS-FALL.2017.8293227).
- WANG Z, XU J, HUANG Z, et al. Road-aided doppler ambiguity resolver for SAR ground moving target in the image domain[J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(10): 1552-1556.
- WANG Z, DU L, MAO J, et al. SAR target detection based on SSD with data augmentation and transfer learning[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(1):150-154.
- WILLIS M J, VON STOSCH M. Simultaneous parameter identification and discrimination of the nonparametric structure of hybrid semi-parametric models[J/OL]. Computers and Chemical Engineering, 2017, 104:366 - 376. <http://www.sciencedirect.com/science/article/pii/S0098135417302041>. DOI: <https://doi.org/10.1016/j.compchemeng.2017.05.005>.
- WILMANSKI M, KREUCHER C, LAUER J. Modern approaches in deep learning for SAR ATR [C/OL]//volume 9843. International Society for Optics and Photonics, 2016. <http://dx.doi.org/10.1117/12.2220290>.
- WINDRIM L, MELKUMYAN A, MURPHY R J, et al. Pretraining for hyperspectral convolutional neural network classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(5):2798-2810.
- WU J, YILDIRIM I, LIM J J, et al. Galileo: Perceiving physical object properties by integrating a physics engine with deep learning[C]//Advances in Neural Information Processing Systems (NIPS). 2015: 127-135.
- WU W, GUO H, LI X. Man-made target detection in urban areas based on a new azimuth stationarity extraction method[J/OL]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2013, 6(3):1138-1146. DOI: [10.1109/JSTARS.2013.2243700](https://doi.org/10.1109/JSTARS.2013.2243700).

- WU W, LI H, ZHANG L, et al. High-resolution PolSAR scene classification with pretrained deep convnets and manifold polarimetric parameters[J/OL]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(10):6159-6168. DOI: [10.1109/TGRS.2018.2833156](https://doi.org/10.1109/TGRS.2018.2833156).
- WURM M, STARK T, ZHU X X, et al. Semantic segmentation of slums in satellite images using transfer learning on fully convolutional neural networks[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150:59-69.
- XIA G, HU J, HU F, et al. AID: A benchmark data set for performance evaluation of aerial scene classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(7):3965-3981.
- XIE H, WANG S, LIU K, et al. Multilayer feature learning for polarimetric synthetic radar data classification[C]//IEEE Geoscience and Remote Sensing Symposium (IGARSS). IEEE, 2014: 2818-2821.
- XIE J, GIRSHICK R, FARHADI A. Unsupervised deep embedding for clustering analysis[C/OL]//International Conference on Machine Learning (ICML): volume 48. 2016: 478-487. <http://proceedings.mlr.press/v48/xieb16.html>.
- XIE Y, FRANZ E, CHU M, et al. Tempogan: A temporally coherent, volumetric gan for super-resolution fluid flow[J/OL]. ACM Transactions on Graphics, 2018, 37(4). <https://doi.org/10.1145/3197517.3201304>.
- YAMAGUCHI Y, YAJIMA Y, YAMADA H. A four-component decomposition of POLSAR images based on the coherency matrix[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2006, 3(3):292-296. DOI: [10.1109/LGRS.2006.869986](https://doi.org/10.1109/LGRS.2006.869986).
- YANG X, YANG J, YAN J, et al. SCRDet: Towards more robust detection for small, cluttered and rotated objects[C]//International Conference on Computer Vision (ICCV). 2019: 8231-8240.
- YANG Y, NEWSAM S D. Bag-of-visual-words and spatial extensions for land-use classification [C/OL]//ACM SIGSPATIAL International Symposium on Advances in Geographic Information Systems (ACM-GIS). 2010: 270-279. <https://doi.org/10.1145/1869790.1869829>.
- YANG Z, DENG J, NALLANATHAN A. Moving target recognition based on transfer learning and three-dimensional over-complete dictionary[J]. IEEE Sensors Journal, 2016, 16(14):5671-5678.
- YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks? [C/OL]//Advances in Neural Information Processing Systems (NIPS). 2014: 3320-3328. <http://papers.nips.cc/paper/5347-how-transferable-are-features-in-deep-neural-networks.pdf>.
- YOU B, YANG J, YIN J, et al. Decomposition of the Kennaugh matrix based on a new norm [J/OL]. IEEE Geoscience and Remote Sensing Letters, 2014, 11(5):1000-1004. DOI: [10.1109/LGRS.2013.2284336](https://doi.org/10.1109/LGRS.2013.2284336).
- YU X, LI Y, JIAO L C. SAR automatic target recognition based on classifiers fusion[C]//International Workshop on Multi-Platform/multi-Sensor Remote Sensing and Mapping. 2011: 1-5.

- YUAN Q, SHEN H, LI T, et al. Deep learning in environmental remote sensing: Achievements and challenges[J/OL]. Remote Sensing of Environment, 2020, 241:111716. <http://www.sciencedirect.com/science/article/pii/S0034425720300857>. DOI: <https://doi.org/10.1016/j.rse.2020.111716>.
- YUAN Y, ZHENG X, LU X. Hyperspectral image superresolution by transfer learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(5): 1963-1974.
- ZHANG H, NASRABADI N M, ZHANG Y, et al. Multi-view automatic target recognition using joint sparse representation[J/OL]. IEEE Transactions on Aerospace and Electronic Systems, 2012, 48(3):2481-2497. DOI: [10.1109/TAES.2012.6237604](https://doi.org/10.1109/TAES.2012.6237604).
- ZHANG L, MA W, ZHANG D. Stacked sparse autoencoder in PolSAR data classification using local spatial information[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(9): 1359-1363. DOI: [10.1109/LGRS.2016.2586109](https://doi.org/10.1109/LGRS.2016.2586109).
- ZHANG Z, WANG H, XU F, et al. Complex-valued convolutional neural network and its application in polarimetric SAR image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(12):7177-7188.
- ZHAO B, HUANG B, ZHONG Y. Transfer learning with fully pretrained deep convolution networks for land-use classification[J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(9):1436-1440.
- ZHAO J, GUO W, LIU B, et al. Preliminary exploration of SAR image land cover classification with noisy labels[C/OL]//IEEE International Geoscience and Remote Sensing Symposium (IGARSS). 2017b: 3274-3277. DOI: [10.1109/IGARSS.2017.8127696](https://doi.org/10.1109/IGARSS.2017.8127696).
- ZHAO J, DATCU M, ZHANG Z, et al. Contrastive-regulated CNN in the complex domain: A method to learn physical scattering signatures from flexible PolSAR images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(12):10116-10135.
- ZHAO J, ZHANG Z, YAO W, et al. OpenSARUrban: A Sentinel-1 SAR image dataset for urban interpretation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13:187-203.
- ZHAO Z, JIAO L, ZHAO J, et al. Discriminant deep belief network for high-resolution SAR image classification[J]. Pattern Recognition, 2017, 61:686-701.
- ZHONG C, MU X, HE X, et al. SAR target image classification based on transfer learning and model compression[J]. IEEE Geoscience and Remote Sensing Letters, 2019, 16(3):412-416.
- ZHOU B, KHOSLA A, LAPEDRIZA A, et al. Learning deep features for discriminative localization [C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2016a.
- ZHOU Y, WANG H, XU F, et al. Polarimetric SAR image classification using deep convolutional

neural networks[J/OL]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(12):1935-1939.

DOI: [10.1109/LGRS.2016.2618840](https://doi.org/10.1109/LGRS.2016.2618840).

ZHU S, LEI B, WU Y. Retrieval of hyperspectral surface reflectance based on machine learning[J].

Remote Sensing, 2018, 10(2):323.

ZHU X X, TUIA D, MOU L, et al. Deep learning in remote sensing: A comprehensive review and list of resources[J/OL]. IEEE Geoscience and Remote Sensing Magazine, 2017, 5(4):8-36. DOI:

[10.1109/MGRS.2017.2762307](https://doi.org/10.1109/MGRS.2017.2762307).

致 谢

终于写到了最后，这也标志着我在中科院电子所五年的博士生活即将进入尾声，这本博士毕业论文也仿佛一部回忆录，记载着五年以来的酸甜苦辣。在博士求学生涯即将结束之际，我想向所有关心、支持、帮助过我的师长和朋友表示衷心的感谢。

首先感谢我的导师雷斌研究员给了我在中科院电子所读博的机会，为我提供了良好的科研环境，给了我充分的自由在科学研究的大海里遨游。雷老师从全局出发的科研观，高屋建瓴的思考，对我后续的论文选题产生了深刻的影响。感谢潘宗序副研究员一直以来对我的指导和帮助，潘老师在科研上带领我入门，言传身教、以身作则，训练我独立科研的能力，让我在问题分析、逻辑思考、论文写作等多个方面都有了很大的提升。感谢仇晓兰研究员在读博期间给予的关心和支持，为我提供了宝贵的国际学术交流机会，仇老师严谨、勤勉的科研态度一直是我学习的榜样。

感谢德国宇航中心的 Mihai Datcu 教授，一直鼓励我大胆表达观点和看法、敢于创新打破常规，做有意义的研究。与 Datcu 教授的合作让我品尝到了做科研的乐趣和激情，愉快又紧张地度过了在德国交流学习的一年，取得了丰硕的成果。感谢教授在我回国后还一直保持电话联系，对我的课题和未来的发展给出了宝贵的指导意见。

感谢课题组的鲍鲜杰老师，实验室的葛蕴萍老师、韩冰老师、谭晓颂老师等为我们提供了科研和生活上的便利和保障。感谢课题组各位同学在读博期间的帮助和陪伴，谢谢陈柯洋和马琳同学对论文提出的修改意见，祝你们早日毕业。感谢 DLR 的姚玮博士、C.O Dumitru 博士、G. Schwarz 博士在数据、实验、论文撰写上提供的帮助和指导意见。感谢玮、兰兰、Yingjie 在生活上的帮助。

感谢博士合唱团的朋友们一起度过了三年快乐的时光，你们是迷雾重重的科研生活中一抹明亮的彩虹。感谢普罗科菲耶夫等大师，你们的音乐在无数个深夜陪伴我做实验写论文，感谢艺术给我的科研带来的灵感。感谢朗朗早日毕业的祝福。

今年是不平凡的一年，感谢一线奋战的抗疫工作人员，也感谢几个月以来居

家隔离的我们，感恩所有人的努力换来了今天的平安。感谢从头到尾在家陪伴我写完毕业论文的父母，谢谢你们多年来的支持和鼓励，一直都是我最坚强的后盾。

曾经山穷水复疑无路，却总是怀揣一颗柳暗花明又一村的希望之心。筚路蓝缕，以启山林，如今终于仿佛是叩开了科研的大门。攀登五年，不知不觉中已翻过一轮山丘，放眼望去，前面还有很长的路，更高的山。

毕业，也是再次出发。

黄钟冷

2020 年 6 月

作者简历及攻读学位期间发表的学术论文与研究成果

作者简历

1994 年 4 月出生于重庆市。

2011 年 9 月至 2015 年 7 月，在北京师范大学信息科学与技术学院进行本科学习，并获理学学士学位。

2015 年 9 月至 2020 年 8 月，在中国科学院电子学研究所攻读博士学位。

2018 年 10 月至 2019 年 9 月，赴德国宇航中心遥感技术研究所联合培养。

已发表 (或正式接收) 的学术论文:

1. **HUANG Z, PAN Z, LEI B**, Transfer Learning with Deep Convolutional Neural Network for SAR Target Classification with Limited Labeled Data. [J] Remote Sensing 2017, 对应论文第二章内容 (SCI 检索, 影响因子: 4.509, 中科院二区, ESI 高被引论文)
2. **HUANG Z, PAN Z, LEI B**, What, Where and How to Transfer in SAR Target Recognition Based on Deep CNNs. [J] IEEE Transactions on Geoscience and Remote Sensing 2020, 对应论文第二章内容 (SCI 检索, 影响因子: 5.855, 中科院二区, TOP 期刊)
3. **HUANG Z, DUMITRU C O, PAN Z, LEI B, DATCU M**, Classification of Large-Scale High-Resolution SAR Images with Deep Transfer Learning. [J] IEEE Geoscience and Remote Sensing Letters 2020, 对应论文第三章内容 (SCI 检索, 影响因子: 3.833, 中科院二区)
4. **HUANG Z, DATCU M, PAN Z, LEI B**, Deep SAR-Net: Learning Objects from Signals. [J] ISPRS Journal of Photogrammetry and Remote Sensing 2020, 对应论文第四章内容 (SCI 检索, 影响因子: 7.319, 中科院一区, TOP 期刊)
5. **HUANG Z, DATCU M, PAN Z, LEI B**. HDEC-TFA: An Unsupervised Learning Approach for Discovering Physical Scattering Properties of Single-polarized SAR Image. [J] IEEE Transactions on Geoscience and Remote Sensing 2020, 对应论文第五章内容 (SCI 检索, 影响因子: 5.855, 中科院二区, TOP 期刊)

6. PAN Z, QIU X, **HUANG Z**, LEI B, Airplane Recognition in TerraSAR-X Images via Scatter Cluster Extraction and Reweighted Sparse Representation. [J] IEEE Geoscience and Remote Sensing Letters 2016 (SCI 检索, 影响因子: 3.534, 中科院二区)
7. **HUANG Z**, DUMITRU C O, PAN Z, LEI B, DATCU M, Can A Deep Network Understand the Land Cover Across Sensors? [C] International Geoscience and Remote Sensing Symposium 2019, 对应论文第二章内容 (EI 检索, 国际会议, 墙报展示)
8. **HUANG Z**, DUMITRU C O, PAN Z, LEI B, DATCU M, A Novel Deep Learning Framework Based on Transfer Learning and Joint Time-Frequency Analysis. [C] TerraSAR-X/TanDEM-X Science Team Meeting 2019, 对应论文第四章内容 (国际会议, 口头报告)
9. **HUANG Z**, DATCU M, Urban Area Analysis in Single-polarized SAR Images Based on Unsupervised Deep Learning. 13th European Conference on Synthetic Aperture Radar 2020, 对应论文第五章内容 (EI 检索, 国际会议, 墙报展示)
10. **HUANG Z**, DATCU M, PAN Z, LEI B, A Hybrid and Explainable Deep Learning Framework for SAR Images. [C] International Geoscience and Remote Sensing Symposium 2020, 对应论文第六章内容 (EI 检索, 国际会议, 口头报告)
11. **HUANG Z**, DATCU M, Deep Learning Specific for SAR Scattering Properties and Image Classification. [C] ESA EO ϕ -week 2020, 对应论文第五章和第六章内容 (国际会议, 口头报告)
12. DATCU M, ANDREI V, DUMITRU C O, **HUANG Z**, SCHWARZ G, ZHAO J, Explainable Deep Learning for SAR Data. [C] ESA EO ϕ -week 2019, 对应论文第三章和第四章内容 (国际会议, 口头报告)

参加的研究项目及获奖情况:

国家自然科学基金项目, 小样本条件下基于深度学习的 SAR 目标识别研究
 十三五全军共用信息系统装备预先研究项目, 小样本机器学习技术在 xxxx 领域的应用研究

获 2017-2018 年度中国科学院大学三好学生